Characterisation of Differential Gene Expression throughout the Life Cycle of Murine Gammaherpesvirus-68

Submitted by Joo Wook Ahn, Wolfson Institute for Biomedical Research, UCL, as part of the requirements for PhD in Virology.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Gammaherpesviruses are associated with a number of diseases including lymphomas and other malignancies. Murine gammaherpesvirus 68 (MHV-68) constitutes the most amenable animal model for this family of pathogens. However experimental characterisation of gammaherpesvirus gene expression, at either the protein or RNA level, lags behind that of other, better-studied alpha and beta herpesviruses.

A differential display system and 2 independent array systems have been developed, with the aim of characterising MHV-68 gene expression profiles and providing experimental supplement to a genome that is chiefly annotated by homology. Furthermore, the 3 technologies are compared in terms of their strengths and weaknesses, with respect to these goals. The MHV-68 transcriptome is presented through a lytic infection *in vitro*, both as quantitative measures of transcript abundance, and as part of a hierarchical cluster analysis. Functional predictions have been made for various genes based on co-expression with better characterised genes. Each gene has also been categorised as being expressed with α-, β- or γ-kinetics by blocking *de novo* protein synthesis and viral DNA replication. This fundamental characterisation furthers the development of this model and provides an experimental basis for continued investigation of gammaherpesvirus pathology.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

4'-S-EtdU	2'-deoxy-5-ethyl-b-4'-thiouridine
A_{260}	Absorbance at 260nm
A_{280}	Absorbance at 280nm
AIDS	Acquired immuno-deficiency syndrome
ATCC	American Type Culture Collection
BAC	Bacterial artificial chromosome
B cell	Lymphoblastoid cell
BCBL	Body cavity based lymphoma
BHK	Baby hamster kidney
BHV-4	Bovine herpesvirus-4
BL	Burkitt's lymphoma
BLAST	Basic local alignment search tool
bp	Base pair
Cab	Calcium binding protein
cdk	Cyclin dependent kinase
cDNA	Complementary DNA
cpe	Cytopathic effect
CX	Cycloheximide
dA	Deoxyadenosine
dATP	Deoxyadenosine 5'-Triphosphate
dC	Deoxycytosine
ddH_2O	Double distilled water
ddPCR	Differential display PCR
DEPC	Diethyl pyrocarbonate
dG	Deoxyguanine

DMEM	Dulbecco's modified Eagle medium
DNA	Deoxyribose nucleic acid
dNTP	Deoxynucleoside triphosphate
dsDNA	Double-stranded DNA
dTMP	Deoxythymidine monophosphate
dT	Deoxythymidine
DTT	Dithiothreitol
E	Early
EBV	Epstein-Barr virus
ECACC	European Collection of Cell Cultures
EDTA	Ethylenediaminetetraacetic acid
EHV-2	Equine herpesvirus 2
FBS	Foetal bovine serum
γ1	Gamma-1
γ2	Gamma-2
GADPH	D-glyceraldehyde -3-phosphate dehydrogenase
GMEM	Glasgow's minimal essential medium
H_2O	Water
HAART	Highly active antiretroviral therapy
HCMV	Human cytomegalovirus
HHV-1	Human herpesvirus-1
HHV-2	Human herpesvirus-2
HHV-3	Human herpesvirus-3
HHV-4	Human herpesvirus-4
HHV-5	Human herpesvirus-5
HHV-6	Human herpesvirus-6
HHV-7	Human herpesvirus-7
HHV-8	Human herpesvirus-8

HPRT	hypoxanthinie phosphoribosyl transferase
HSV	Herpes simplex virus
HVS	Herpesvirus saimiri
ICP	Infected cell protein
IE	Immediate-early
IM	Infectious mononucleosis
i.n.	Intranasal
i.p.	Intraperitoneal
IPTG	Isopropyl β-D-thiogalactopyranoside
kb	Kilobase
kbp	Kilobase pairs
KS	Kaposi's sarcoma
KSHV	Kaposi's sarcoma-associated herpesvirus
L	Late
L-DNA	Central unique segment
lacZ	Beta galactosidase
LANA	Latency-associated nuclear antigen
LAT	Latency-associated transcript
MHV-68	Murine gammaherpesvirus-68
MMLV	Mouse Moloney murine leukemia virus
moi	Multiplicity of infection
MOPS	3-[N-morpholino]propanesulfonic acid
NaOAc	Sodium acetate
NCBI	National Centre for Biotechnology Information
NCS	Newborn calf serum
NIH	National Institutes of Health
NTP	Ribonucleoside triphosphate
o/n	Overnight

ORF	Open reading frame
PAA	Phosphonoacetic acid
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEL	Primary effusion lymphomas
pfu	Plaque-forming units
pi	Post-infection
poly-A	Poly-adenosine
RNA	Ribonucleic acid
RPM	Rotations per minute
rRNA	Ribosomal RNA
RT	Reverse transcriptase
RT-PCR	Reverse transcription and PCR
RTase	Reverse transcriptase
SCID	Severe combined immune deficient
SDS	Sodium dodecyl sulphate
SSC	Saline-sodium citrate buffer
ssDNA	Single-stranded DNA
SSPE	Saline-Sodium Phosphate-EDTA Buffer
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TCA	Trichloroacetic acid
$TCID_{50}$	Tissue culture infective dose 50
TE buffer	Tris-EDTA buffer
tk	Thymidine kinase
TMV	Tobacco mosaic virus

TPA	12-O-tetradecanoylphorbol-13-acetate
Tris	Tris-(hydroxymethyl)aminomethane
tRNA	Transfer RNA
U_L	Unique long region
U_S	Unique short region
UV	Ultra-violet light
VEGF	Vascular endothelial growth factor
VHS	Virion-associated host shutoff
VP	Viral protein
VZV	Varicella zoster virus
wt	Wild type
w/v	Weight to volume
w/w	Weight to weight
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

ACKNOWLEDGEMENTS

To Jette, my parents, my sisters and my friends.

To everyone who has helped me in this PhD, including Helen and Elaine; Paul Kellam, Pete Corish, Marcus Harrison, Stacy Efstathiou, James Stewart, Marketa Zvelebil, Stephen Henderson, Neale Foxwell, Dave Goodwin, James Leiper.

Lastly and not leastly to Dagmar and Ken.

1 Introduction

1.1 Family *Herpesviridae*

1.1.1 History of the *Herpesviridae* Family

The ancient Greeks first coined the term 'herpes', deriving it from their word *herepein*, which means 'to creep'. There are numerous references to herpes-like diseases dating from ancient Sumeria (tablet dated 3rd millennium BC) and ancient Egypt (Ebers papyrus dated 1552 BC; reviewed in Roizman, 2001b) In 1873 AD, herpes disease was first demonstrated to be transmissible between two humans by Vidal, and in 1924, Grunter showed conclusively that herpes disease was the result of an infectious agent via an extensive series of animal to animal propagation experiments (reviewed in Whitley, 1998). In 1939, Burnet and Williams pronounced that herpes simplex virus (HSV) infections were lifelong, and that the virus remained latent until certain stimuli recalled the virus into activity, resulting in a lesion (Burnet, 1939). This was the first suggestion that herpesvirus infections become latent in their host, and can reactivate.

There have since been huge advances both in our knowledge of herpesviruses and in the experimental tools available to virologists. Indeed, the development of herpesvirus biology has often been punctuated by the development of new technologies. The discovery that HSV could be cultivated in the chorioallantoic membrane of chicken eggs gave scientists the option to move away from the animal models that had been the primary experimental system before (Andrews, 1930). In particular, this method of cultivating the virus made possible various immunological studies of herpesviruses, through neutralisation and complement fixation reactions (Burnet, 1937).

The development of *in vitro* tissue culture technology in the 1950's led to the discovery of varicella zoster virus (VZV; Weller, 1952) and human cytomegalovirus (HCMV; Smith, 1956; Rowe, 1956). Advancement of this technique, to allow cultivation of B lymphocytes and lymphoblastoid tumour cells, led to the observation of virus particles in lymphoblastic tumour cells (Epstein, 1957). This virus was characterised as Epstein-Barr virus (EBV) and was the first human gammaherpesvirus to be discovered (Epstein, 1965). The introduction of electron microscopy to virology, around this same period, was instrumental for developing the field as viruses could be visualised for the first time (Brenner, 1959; Wildy, 1963).

Further development of tissue culture techniques, to allow cultivation of T lymphocytes, led to the discovery of yet more human herpesviruses by the 1990's: herpesvirus-6 (HHV-6) and -7 (HHV-7; Salahuddin, 1986; Frenkel, 1990). Herpesvirus-8 (HHV-8) or Kaposi's sarcoma-associated herpesvirus (KSHV) was discovered by representational differential analysis (Chang, 1994) and was the second human gammaherpesvirus.

Degenerate polymerase chain reaction (PCR) has been used to detect novel gammaherpesviruses in pigs (Chmielewicz, 2003) and baboons (Whitby, 2003). Chimpanzees are known to harbour 3 distinct lineages of gammaherpesviruses and there are 2 distinct lineages in macaques (Rose, 1997; Lacoste, 2000; Lacoste, 2001), and so one could predict that a third human gammaherpesvirus awaits discovery.

Recently, array technologies for the efficient analysis of gene expression at the genome level have been developed. Coupled with the whole genome sequence data that is available for many viruses, these technologies could provide very useful for the study of herpesviruses. Also, as the regulation of gene expression has been suggested as the key factor in determining an organism's behaviour, rather than its genomic sequence (Levine, 2003), arrays present a powerful tool with which to examine this hypothesis. It remains to be seen how much this particular technological advance will influence herpesvirus biology.

1.1.2 Structural Definition of Herpesviruses

Viruses were initially characterised by their small size as assessed by filterability, but as the structure and composition of viruses became more apparent, viruses were grouped according to shared virion properties (Andrewes, 1954). The herpesviruses were among the first taxonomic groups to be constructed, based chiefly on the formation of Cowdry type-A intranuclear inclusion bodies following infection by these viruses (Andrewes, 1954; Andrewes, 1961). The chemical composition of viruses was also used to characterise them. However, the development of electron microscopy, and in particular negative staining, allowed structural definitions of viruses to become commonplace (Brenner, 1959; Wildy, 1963). Illustrations of herpesvirus structure are shown in Figure 1.1.

Herpesviruses are enveloped viruses with a double-stranded deoxyribose nucleic acid (dsDNA) genome. The envelope is derived from host lipids, from which extend glycoprotein spikes encoded by the viral genome. While the combination of glycoproteins making up these spikes varies greatly between members of the *Herpesviridae*, their function is generally conserved. The glycoprotein spikes

mediate attachment and entry of the virus into cells, cell-to-cell spread of infection, and also influence tissue tropism and host range (reviewed in Cole, 2003).

HSV-1's tegument consists of four major structural proteins (Spear, 1972; Honess, 1973) and four other gene products. These tegument proteins are the first to be exposed to the intracellular environment of a host cell and are thought to provide critical viral functions in the time between viral penetration of the cell and the synthesis of viral immediate-early (IE) proteins, such as shutoff of host cell protein synthesis (Fenwick, 1978; Smibert, 1992) and immediate-early gene transactivation (Campbell, 1984).

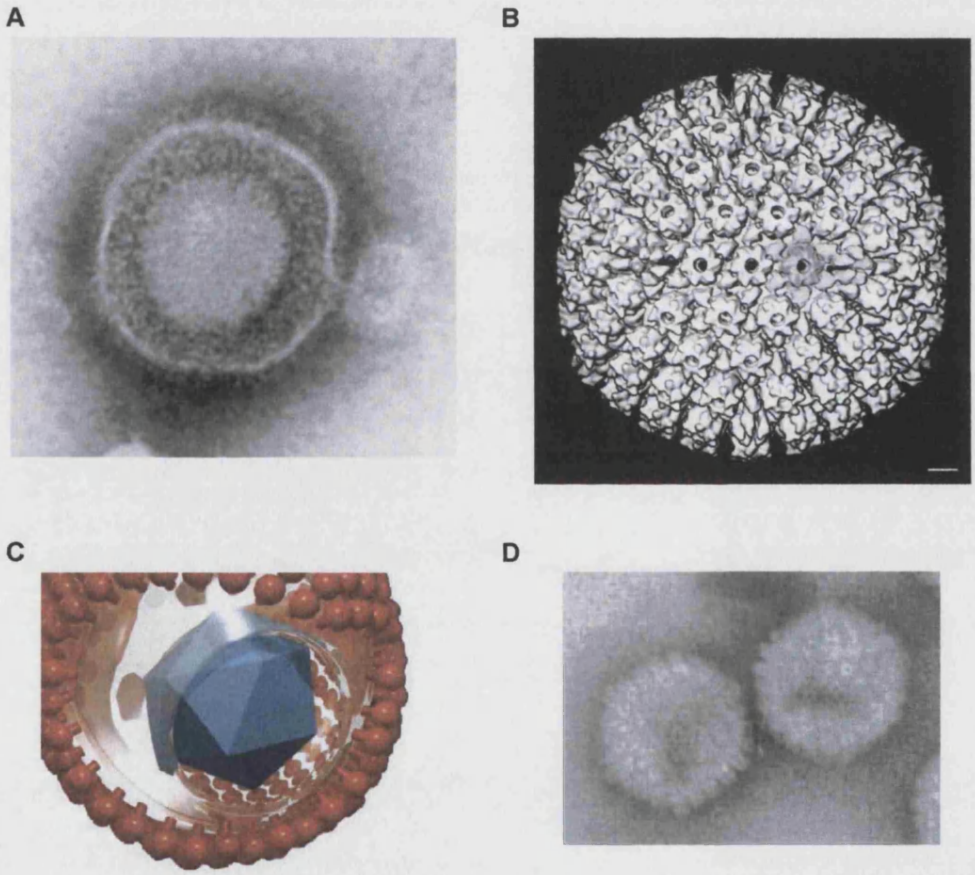


Figure 1.1. Herpesvirus structure.
A. Electron micrograph of HSV-1 virion (taken from Sander, 2003). **B.** 3-D computer reconstruction of HSV-1 outer capsid from cryo-electron micrographs, bar =100 Å (adapted from Cheng, 2002). **C.** 3-D computer representation of herpesvirus structure (Reschke, 1995). Glycoproteins in the envelope are shown in red, the envelope itself and tegument are shown as transparent, and the capsid is shown in blue. **D.** Electron micrograph of mature isocahedral HSV capsid (adapted from Newcomb, 1999).

The HSV-1 capsid is isocahedral (T=16), 100-110nm in diameter, consisting of 150 hexons and 12 pentons, each with a channel down the long axis (Wildy, 1963; Wu, 2000a). Inside the capsid lies the nucleoprotein core, which contains the linear, dsDNA genome in the form of a torus. Herpesvirus genomes range from about 100-235 kilobases pairs (kbp). The HSV-1 genome consists of 2 regions known as the unique long (U_L) and unique short (U_S) regions. These are bound by repeat regions, which form the internal and terminal repeats. Genome sizes can vary within a species as these repeat regions contain varying numbers of repeats. As a result genomes can range in size by more than 10kilobases (kb) within a species. The family can be divided into 6 groups based on their sequence arrangements (Roizman, 1996). Between species, gene order is generally maintained within large sections of the genome, with varying degrees of homology between the genes themselves. Homologous gene sequences between members of the herpesvirus family can be used to construct evolutionary trees (Gompels, 1988; Montague, 2000).

1.1.3 Herpesvirus Taxonomy

The *Herpesviridae* are divided into three subfamilies: *alphaherpesvirinae*, *betaherpesvirinae* and *gammaherpesvirinae*, and there are also 61 unclassified herpesviruses currently listed in the ICTV database (as of 11/2003; Buchen-Osmond, 2002). The taxonomic structure of the herpesvirus family and examples of human herpesviruses are listed in Table 1.1. Currently, viral taxonomy considers several factors including morphology, physicochemical and physical properties, protein, lipid and carbohydrate composition, antigenic properties and biological properties (Regenmortel, 2000). In particular, sequence homology-based classifications of herpesviruses are gaining prominence (McGeoch, 1995; McGeoch, 2000).

Genetic methods of classification have the advantage that they are inherently more indicative of evolutionary relationships. These methods (in particular the sequence comparison algorithms) are becoming more sophisticated, and multiple gene sequence comparisons are now commonplace (Karlin, 1994; Montague, 2000; Alba, 2001). A recent attempt at classifying the family using the sequence of the conserved glycoprotein B gene is shown in Figure 1.2 (Sharp, 2002).

1.1.3.a Subfamily Gammaherpesvirinae

The 22 members of the *gammaherpesvirinae* subfamily are usually associated with a latent infection of lymphoblastoid cells, and usually with a lytic infection of epithelial or fibroblast cells. A further 6 herpesviruses have been tentatively

classified as gammaherpesviruses, pending further characterisation of these species (as of 11/2003; Buchen-Osmond, 2002). The subfamily is divided into two genera, which are known as the gamma-1 (γ1) subgroup or lymphocryptoviruses, and gamma-2 (γ2) subgroup or rhadinoviruses (Roizman, 1996). There are 2 known human gammaherpesviruses, EBV and HHV-8, which are part of the γ1 and γ2 subgroups, respectively. The two gammaherpesvirus genera are well defined by genome organisation and gene content. For example, the rhadinoviruses have a central unique segment of 110-141kb with a low G+C ratio (known as the L-deoxyribose nucleic acid (L-DNA)) that is flanked by repeated regions of high G+C content.

Some gammaherpesviruses viruses are difficult to assign to either one of the two genera. For example, equine herpesvirus 2 (EHV-2) is clearly a γ2 herpesvirus based on gene content, but possesses an atypical genomic structure and G+C distribution (Telford, 1995). As 4 genera exist for alphaherpesviruses, and 3 for betaherpesviruses, it remains to be seen whether further genera will be assigned to the gammaherpesviruses.

Table 1.1. Taxonomic Structure of the *Herpesviridae* Family.

Taxonomic Level	**Group Name**	**Human herpesviruses**
Subfamily	*Alphaherpesvirinae*	
Genus	*Simplexvirus*	Herpes simplex virus-1 (HHV-1) Herpes simples virus-2 (HHV-2)
Genus	*Varicellovirus*	Varicella zoster virus (HHV-3)
Genus	"Marek's disease-like viruses"	
Genus	"Infectious laryngo-tracheitis-like viruses"	
Subfamily	*Betaherpesvirinae*	
Genus	*Cytomegalovirus*	Human cytomegalovirus (HHV-5)
Genus	*Muromegalovirus*	
Genus	*Roseolovirus*	Human herpesvirus-6 (HHV-6) Human herpesvirus-7 (HHV-7)
Subfamily	*Gammaherpesvirinae*	
Genus	*Lymphocryptovirus*	Epstein-Barr virus (HHV-4)
Genus	*Rhadinovirus*	Human herpesvirus-8 (HHV-8)

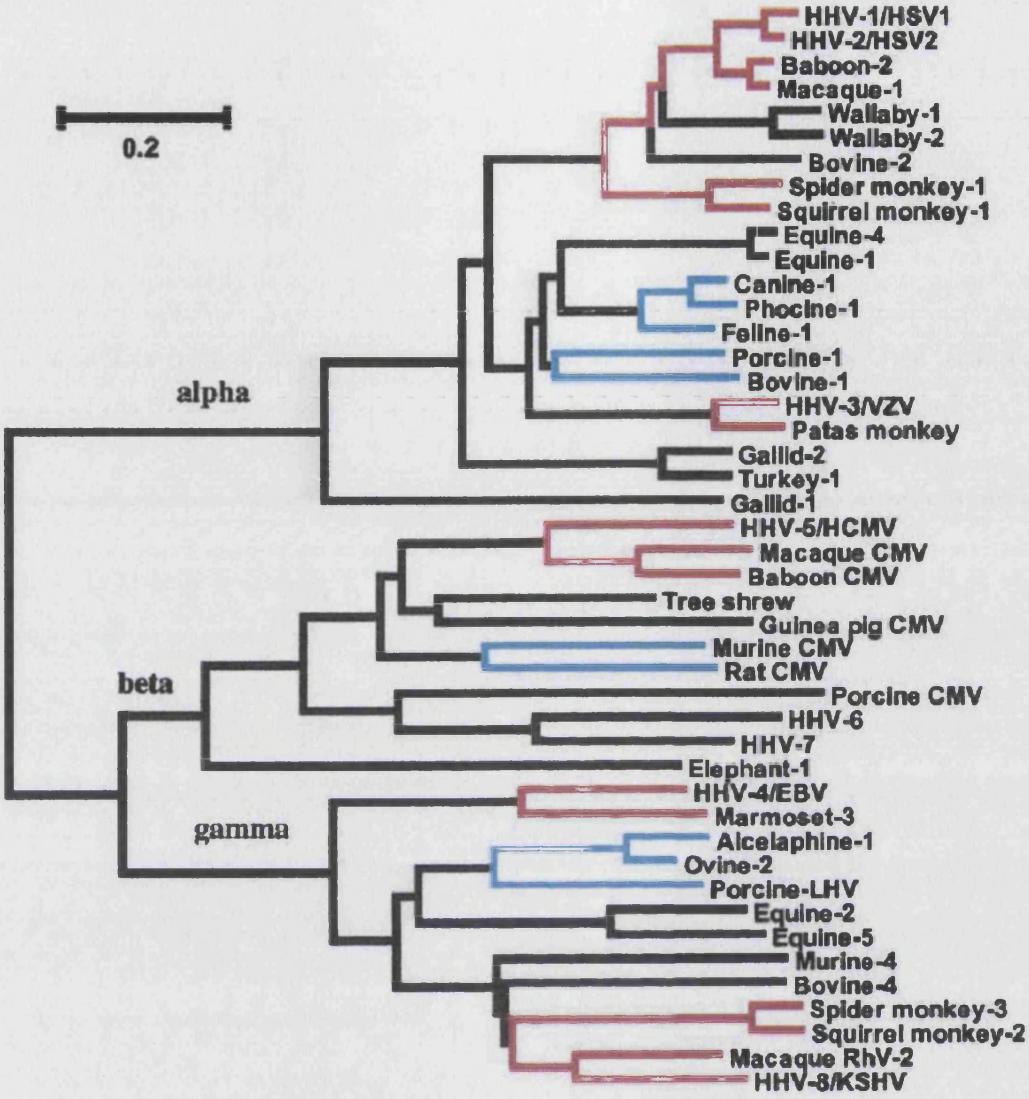


Figure 1.2. Phylogeny of herpesviruses, derived from comparisons of glycoprotein B sequences of mammalian-host herpesviruses.

Branching patterns consistent with virus-host co-speciation are highlighted for viruses from primates (red) and from carnivores, artiodactyls, and rodents (blue). The scale bar indicates 20% sequence divergence. The eight human viruses are labelled HHV; for other viruses, the species of origin is indicated. MHV-68 is denoted as Murine-4 (from Sharp, 2002).

1.1.3.b Genus Rhadinovirus

The only human rhadinovirus known to date is HHV-8, but rhadinoviruses have been isolated from a variety of organisms. Murine gammaherpesvirus-68 (MHV-68) is a recently discovered rhadinovirus that appears to be slightly divergent from other gammaherpesviruses. This is demonstrated in its glycoprotein B sequence homology to other gammaherpesviruses as shown in Figure 1.2, its major capsid protein sequence (Trus, 2001), and perhaps most clearly in its biological properties.

Although sequence analysis of the MHV-68 genome shows it to be a γ2-herpesvirus (Virgin, 1997), it possesses some biological properties that are more typical of alphaherpesviruses: It has a rapid replication cycle and can infect a broad range of cell lines *in vitro* (Svobodova, 1982a; Mistrikova, 1985). However MHV-68 also has a latent phase of infection in lymphoblastoid cells (B cells) (Sunil-Chandra, 1992b) and has been associated with tumours (Sunil-Chandra, 1994), which are properties characteristic of gammaherpesviruses.

MHV-68 has sparked interest as a model system for the gammaherpesviruses due to the practical advantages of working with this virus (detailed in 1.2.2.c). It has also proved to be an interesting subject for study in its own right. For example, MHV-68 is the only virus that has been shown to encode an abundantly secreted chemokine binding protein (van Berkel, 1999; Parry, 2000; Alexander, 2002).

1.1.4 Disease and Pathogenesis of Herpesviruses

The main diseases associated with primary infection and reactivation of the human herpesviruses are shown in Table 1.2. Herpesviruses are associated with a wide range of diseases and are generally widely prevalent in their host population: HSV-1, and HCMV can show seroprevalences above 90% in the Western world (de Jong, 1998; Smith, 2002). HHV-8 is an exception as its seroprevalence is far lower at 2-4% in Europe, Southeast Asia and the Caribbean. However in areas of sub-Saharan Africa where it is endemic, seroprevalence reaches over 50% (Olsen, 1998; Ablashi, 1999; DeSantis, 2002). While many herpesviruses are widespread in their host populations, herpesvirus infections are generally asymptomatic. However, these viruses do have the potential to cause life-threatening infections under a certain conditions.

Most of the historical references to herpesviruses relate to herpes labialis (cold sores). Herpes labialis is generally a result of HSV reactivating from the ganglia of peripheral neurones and migrating down axons to release progeny virus, usually in the region of the lips. Occasionally, primary HSV infection of newborn children can cause congenital herpes disease, which can be life threatening when it results in an encephalitis or a disseminated herpes infection. HCMV is the leading cause of congenital viral infections now that rubella is controlled by vaccination, and this type of HCMV infection can lead to mental retardation and deafness (Revello, 2002).

Many human herpesviruses have gained prominence due to their association with transplant immunosuppression and acquired immuno-deficiency syndrome (AIDS): HCMV and HHV-6 with solid organ transplants (Rubin, 2001), bone marrow transplants and stem cell transplants (Boutolleau, 2003), and HSV-1, HHV-6, and

HCMV with AIDS (Griffiths, 2002). The immunocompromised nature of these patients allows reactivating herpesvirus infections to progress relatively unchecked. The same occurs in cancer patients undergoing therapies that also result in an immunocompromised state. Other diseases that may have a herpesvirus component include atherosclerosis and multiple sclerosis (Alber, 2000; Simmons, 2001).

Gammaherpesviruses are associated with various malignancies. EBV is associated with Burkitt's lymphoma (BL), Hodgkin's disease, nasopharyngeal carcinoma, and other lymphoproliferative diseases. Burkitt's lymphoma is an aggressive B cell tumour, the endemic form of which accounts for more than half of all childhood cancers in equatorial Africa (Magrath, 1991). There is a morphologically identical sporadic form that is more common in the USA and Europe, and BL is also common among HIV-positive patients, which represents a third form. All forms of BL show a chromosomal translocation of *c-myc*, resulting in increased levels of this transcription factor (Taub, 1982; Dalla-Favera, 1982).

Although EBV was first isolated from a BL cell line, and there is a clear epidemiological link between EBV and BL, it has not been possible to prove causation or to elucidate the mechanisms by which the virus initiates transformation. In particular while over 90% of endemic BL cases are associated with EBV infection, sporadic and HIV-related forms show 10-85% and 30-80% association, respectively (Gutierrez, 1992; Davi, 1998), and this has been a source of uncertainty in determining the aetiological role of EBV.

HHV-8 is associated with Kaposi's sarcoma (KS), multicentric Castleman's disease and primary effusion lymphoma. KS is the most common cancer in AIDS patients, and at the onset of the AIDS epidemic, homosexual men in the USA had a 50% lifetime chance of developing KS (Katz, 1994). The development and introduction of highly active antiretroviral therapy (HAART) in Western countries has successfully reduced this figure, but less than 1% of AIDS patients receive HAART in sub-Saharan Africa, where 70% of all global cases occur (Harries, 2002).

There are various other forms of KS such as classic KS, which occurs predominantly in elderly male patients of Southern European ancestry. A high frequency is also seen in Israel and other Middle Eastern countries (Ablashi, 2002). Endemic KS occurs in some parts of sub-Saharan Africa, where it is usually a childhood cancer, increasing in incidence since the onset of the AIDS epidemic. It is generally a more aggressive disease than classic KS, and children infected with it tend to develop lymphadenopathic tumours, which are often fatal (Ziegler, 1996).

Table 1.2. Examples of the more common diseases associated with the human herpesviruses.

Virus	Some common diseases associated with primary infection	Some common diseases associated with reactivation
HSV-1	Herpes labialis	Herpes labialis
HSV-2	Herpes genitalis Various localised and disseminated diseases resulting from Congenital infection	Herpes genitalis
VZV	Varicella	Herpes zoster
HCMV	Various CNS and systemtic diseases resulting from Congenital infection	Usually subclinical except in immunocompromised patients
HHV-6	Febrile illnesses including exanthem subitum	Usually subclinical except in immunocompromised patients
HHV-7	Febrile illnesses including exanthem subitum	Usually subclinical except in immunocompromised patients
EBV	Infectious mononucleosis	Burkitt's lymphoma Hodgkin's disease
HHV-8	Usually subclinical	Kaposi's sarcoma Primary effusion lymphoma Multicentric Castlemans disease

1.2 Murine Gammaherpesvirus – 68

1.2.1 Isolation and Initial Characterisation of MHV-68

In November 1976, Blaškovič, Stančeková, Svobodová and Mistríková of the J. A. Komensky University in Bratislava performed a field study on the ecology of arboviruses in Slovakia (then Czechoslovakia). A number of small rodents were trapped and as part of the survey, were examined for the presence of viruses. From the 19 samples collected from bank voles (*Clethrionomys glariolus*) and wood mice (*Apodemus flavicollis*), 5 infectious viral strains were isolated and designated isolates 60, 68, 72, 76 and 78 (Blaskovic, 1980).

Baby hamster kidney cells (BHK) infected with the viral isolates became rounded up with swollen nuclei, which showed hourglass shaped intranuclear inclusions with marginalisation of host chromatin (Blaskovic, 1980; Ciampor, 1981). These features suggested that the infectious pathogens were part of the herpesvirus family.

These putative herpesviruses successfully infected 16 cell lines *in vitro*, including those of mouse, rabbit and human origin. Rabbit embryo fibroblast and BHK cell lines produced the highest titres of infective virus at 10^7 and 10^6 tissue culture infective dose 50 ($TCID_{50}$), respectively, while the human epithelial cell lines produced titres of 10^2-10^3 $TCID_{50}$. Based on this broad range of host cell lines being permissive for infection, and the observation of rapid growth curves, the isolates were classified as alphaherpesviruses (Svobodova, 1982a).

The murine herpesvirus isolates were found to be antigenically related by complement fixation and virus neutralisation assays (Svobodova, 1982b), and therefore a representative murine herpesvirus was selected for further study. Murine herpesvirus isolate 68 was chosen, as it was the most practical to work with, as discussed later.

A series of experiments were performed to confirm this putative classification, culminating with the sequencing of the virus' genome (Blaskovic, 1984; Rajcani, 1986; Efstathiou, 1990; Virgin, 1997). While the virus was confirmed to be a herpesvirus, it was found from its deoxyribose nucleic acid (DNA) sequence to be part of the gammaherpesvirus subfamily, albeit an atypical member compared to the human herpesviruses.

1.2.2 MHV-68 as a Model System for the Gammaherpesviruses

1.2.2.a *In Vitro* Model Systems for Gammaherpesviruses

Much of the knowledge on gammaherpesviruses stems from work on EBV, which is the most extensively studied gammaherpesvirus to date. There is also a rapidly growing base of knowledge for HHV-8, which has received a lot of attention since its discovery (Chang, 1994). While some facets of gammaherpesvirus biology are relatively well researched, such as latency, there is still much that remains unknown. Even the site of EBV primary productive infection is unclear. Currently there is evidence for epithelial cells and B cells being the site of primary infection (Allday, 1988; Faulkner, 1999).

Gammaherpesvirus research has been hindered by the absence of a well-established model system for lytic infection. EBV and HHV-8 have narrow cell tropisms, which is typical for gammaherpesviruses, and there is a lack of permissive cell lines for studying these viruses (Blackbourn, 2000; Bornkamm, 2001). Lytic gammaherpesviruses infections have instead been studied by inducing reactivation of latently infected cell lines, usually with 12-O-tetradecanoylphorbol-13-acetate (TPA) or sodium butyrate. These cell lines are derived from

gammaherpesvirus-associated tumours and therefore may not be an accurate model of latency (Cen, 1993). Furthermore, induced reactivation does not necessarily show the same mechanisms as natural reactivation *in vivo*, especially considering that the induction agents induce some viral promoters directly (Farrell, 1983; Faulkner, 1999). Therefore, studies using these cell lines may not be accurate representations of primary productive infection or reactivation *in vivo*. However these tumour-derived cell lines have proved useful as BCBL-1 cells (body cavity based lymphoma; derived from HHV-8 infected primary effusion lymphoma cells) are one of the only readily available sources of HHV-8 (Renne, 1996). Also, Burkitt's and Hodgkin's lymphoma-derived cell lines, which harbour latent EBV, have helped characterise the various latency programmes of EBV (Crawford, 2001). Attempts to study EBV by infecting primary human cells have has limited success (Robinson, 1975; Savard, 2000). Cultivating HHV-8 *in vitro* has also proved problematical as those cell lines that can be infected tend only to be partially permissive for HHV-8 lytic infection (Renne, 1998). HVS subgroup C can transform human lymphocytes *in vitro* (as well as rabbit, rhesus monkey and it's natural host the spider monkey), and there have also been many studies of HVS transformed cells (Hall, 2000; Whitehouse, 2003).

However, the latent life cycle of EBV is probably the best-understood example of herpesvirus latency. These studies have been aided by the fact the EBV will infect B cells *in vitro*, and will efficiently activate and transform them. This has served as a powerful model for EBV latency and initiation of lymphoproliferative disease (Sugden, 1994). However, at the same time, this property is also the weakness in the model, as *in vivo*, EBV infection of B cells rarely leads to lymphoproliferative disease.

1.2.2.b *In Vivo* Model Systems for Gammaherpesviruses

Attempts have been made to model EBV in non-human primates, but have had limited success due to cross-reacting antibodies against simian herpesviruses. Some New World monkeys have proved susceptible to EBV infection, such as the cotton marmoset (*Callithrix jacchus*) and the cottontop tamarin (*Saguinus oedipus oedipus*) (Epstein, 1985; Cox, 1996). However these model systems are limited by differing outcomes of disease, and the practical limitations of working with endangered species.

Severe combined immune deficient (SCID) mice can also be used to model the human immune system by inoculating them with human peripheral blood mononuclear cells (PBMC). When EBV seropositive blood donors are the source of

PBMC, 45% of SCID mice are observed to develop lymphoma (Rochford, 1995). This system has been used to define the tumour types that can arise following EBV infection. There have also been efforts to model HHV-6 infection in this system (Dittmer, 1999; Aoki, 1999; Lane, 2002). However there is a relatively high degree of variability in this system as 3-9 month old SCID mice can develop a limited number of B and T cells, and up to 15% of SCID mice may spontaneously develop T cell lymphomas (Johannessen, 1999). Furthermore, differing methods are used to pre-condition mice to accept human cells, which introduces further variability. This system has yet to be standardised and therefore cross-comparison of data can be difficult.

1.2.2.c MHV-68 as a Model System for Gammaherpesviruses

Unlike other gammaherpesviruses, MHV-68 readily infects a broad range of cell lines, including primate and rodent-derived lines (Mistrikova, 2000). Therefore, MHV-68 can be used to infect well-defined tissue culture systems without a need for specialised conditions. Combined with its rapid replication cycle, it represents a very practical *in vitro* model system for the gammaherpesviruses. Furthermore, being a natural pathogen of rodents, the virus can be used to infect laboratory mice, which presents a natural small animal system in which to model gammaherpesviruses *in vivo.*

Recently, a bacterial artificial chromosome (BAC) based mutagenesis system for MHV-68 has been introduced, allowing relatively simple creation of mutant strains (Adler, 2000). This provides a reliable method for manipulating the virus. As there are a wide variety of cell lines and mouse strains are available, defined changes to the virus' environment can also be made. Therefore, both the pathogen and its environment can be controlled and modified. Such a model system has many advantages for gammaherpesvirus research, and in particular, human gammaherpesvirus research can benefit greatly as many of the restrictions imposed on experiments with HHV-8 and EBV are obviated. However it remains to be seen exactly how similar the biology of MHV-68 and the human gammaherpesviruses are.

1.2.3 General Life Cycle of Herpesviruses

The life cycle of MHV-68 is typical of a herpesvirus in that it consists of attachment and penetration of the virus into the host cell, translocation of the viral genome into the host nucleus, transcription of viral genes, replication of viral DNA, and the assembly and release of progeny virus. A generalised herpesvirus lytic life cycle is shown in Figure 1.3.

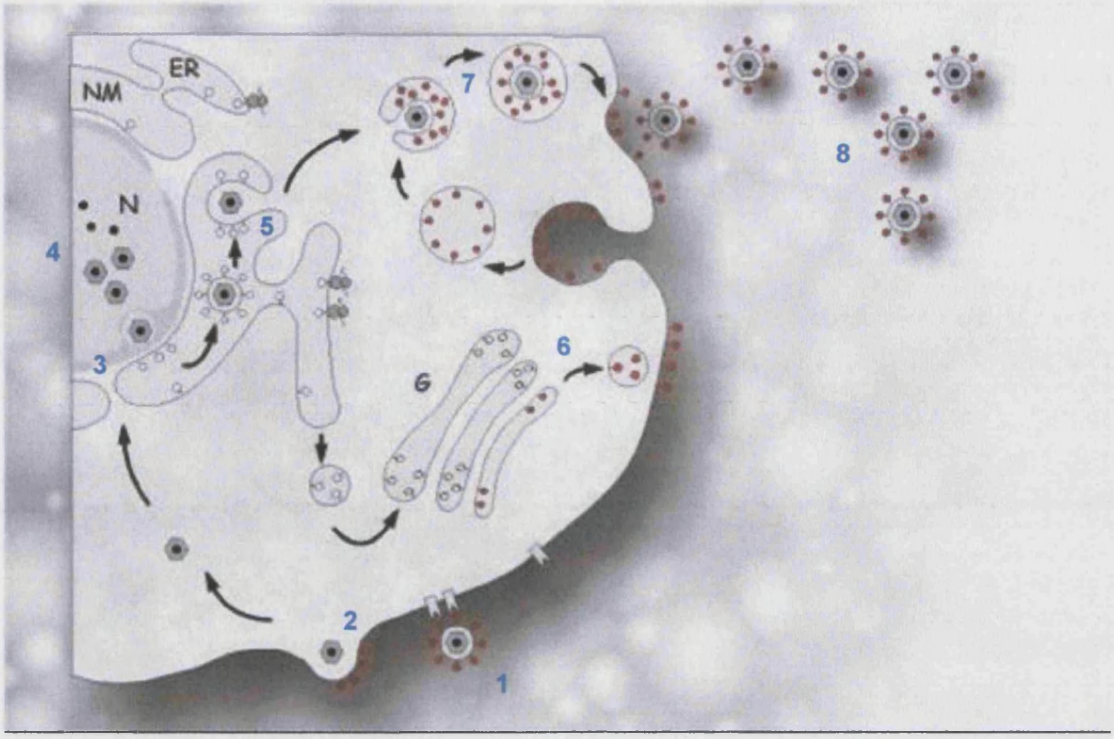


Figure 1.3. Generalised representation of a herpesvirus lytic life cycle.
Numbers represent the various stages as follows: 1 - attachment, 2 - entry into cell, 3 - entry into nucleus, 4 - transcription, replication and assembly of new virus particles, 5 - exit from nucleus, 6 - assembly of glycoproteins to cell surface, 7 - de-envelopment/re-envelopment, 8 - release of progeny virus. Abbreviations: N - nucleus, NM - nuclear membrane, ER - endoplasmic reticulum, G - golgi apparatus (Adapted from Reschke, 1995)

1.2.3.a Attachment and Penetration

Much of the knowledge of the herpesvirus life cycle is derived from work on HSV-1. Attachment and penetration of HSV-1 relies on sequential interactions between specific viral glycoproteins and host cell surface receptors, initially between viral glycoprotein B or C and cell surface heparin sulphate (Johnson, 1984; Addison, 1984; Epstein, 1984; WuDunn, 1989; Shukla, 2001). Heparin sulphate is also involved in the attachment of HMCV, HHV-7 and HHV-8 as well, but EBV attachment appears to be independent of heparin sulphate (Compton, 1993; Secchiero, 1997; Speck, 2000; Akula, 2001). After this initial step, viral glycoprotein D interacts with further cell surface molecules, which is thought to instigate virus-cell fusion (Ligas, 1988). The action of other viral glycoproteins is required to complete adsorption of the virus (Sarmiento, 1979; Campadelli-Fiume, 1990; Hutchinson, 1992; Forrester, 1992; Dingwell, 1994).

Following entry, the nucleocapsid is transported to the nuclear pore complexes (Miyamoto, 1971). Studies with HSV-1 have shown that around 70% of bound virus penetrates with a half-time of around 8min (Sodeik, 1997). Over 60% of these capsids are delivered to the nuclear pore complexes within 4h of entry, with the majority of the tegument being left behind. The capsids travel along the microtubule network and the viral genome is released into the nucleus (Sodeik, 2000; Ploubidou, 2001).

The viral DNA is accompanied by a tegument protein VP16 (also known as α-TIF, ICP25; the product of gene U_L48), which enhances immediate early viral transcription via cellular transcription factors (Campbell, 1984; Flint, 1997). Another tegument protein, the virion-associated host shutoff protein (vhs; product of gene U_L41) remains in the cytoplasm where it causes the disaggregation of polyribosomes and degradation of cellular and viral RNA (Schek, 1985; Smibert, 1992).

1.2.3.b Transcription

Transcription of the viral genome then occurs in the nucleus in 3 distinct stages, known as immediate-early (IE), early (E) and late (L) (Jones, 1979). These designations refer back to studies of bacteriophage gene expression, and a modern terminology of alpha (α), beta (β) and gamma (γ) kinetic classes of transcription has been suggested to help avoid confusion between 2 different transcription systems. Much of the work on herpesvirus transcription has centred on HSV-1, as it has been the most practical virus to work with.

α genes are transcribed without the need for any viral protein synthesis and are involved in priming the cell for further viral gene expression. Of the close to 100 transcripts encoded by HSV-1, only 5 are expressed as α genes. The HSV-1 tegument protein, α-TIF, also enters the nucleus with the viral DNA and interacts with the host's transcription factors (Batterson, 1983; Xiao, 1990). The resulting complex binds to α-gene promoters in the viral genome and initiates transcription of α genes by the host's RNA polymerase II (Campbell, 1984; Preston, 1988).

Three of the proteins encoded by HSV-1's α genes then form further nuclear complexes and are involved in initiating β transcription (genes α0, α4 and U_L54 encoding proteins ICP0, ICP4 and ICP27, respectively). The α gene α22 (encoding ICP22) is essential in most but not all cell lines and is involved in initiation of γ gene transcription (Post, 1981; Long, 1999). The last α gene, U_S12 (encoding ICP47), is

involved in modulating the host's immune response by blocking antigen presentation (York, 1994).

β gene transcription is defined by it's requirement for viral protein synthesis, as their transcription requires redirection of the host's transcriptional machinery by the α gene products (Honess, 1974). The β gene products are primarily involved in viral DNA replication and as soon as sufficient levels of β proteins are produced, viral DNA replication commences. β gene promoters are unlike α gene promoters in that they are diverse and do not share common elements (Roizman, 2001a). Therefore, the expression of β genes cannot be predicted from their promoter sequences and their expression profiles can vary greatly. In HSV-1, U_L39, U_L40 and U_L29 (ribonucleotide reductase, large and small subunits, and major DNA binding protein, respectively) are amongst the first β genes to be expressed (Conley, 1981; Huszar, 1981), while others such as U_L23 and U_L30 (thymidine kinase and DNA polymerase, respectively) are expressed later (Klemperer, 1967; Purifoy, 1977).

The γ genes tend to encode structural components of virus particles. They are further divided into γ1 and γ2 classes, which are also referred to as early or leaky late and true late classes, respectively (Jones, 1979). Transcription of γ1 genes commences around the same time as the β genes but their expression levels only peak after DNA replication has been initiated (e.g. HSV-1 U_L27 and U_s6, glycoproteins B and D, respectively). γ2 genes are only expressed following the onset of viral DNA replication (e.g. HSV-1 U_L44, glycoprotein C). Similar to β genes, the promoter sequences of γ genes do not share any common elements, which explains the diversity of their expression patterns (Roizman, 2001a).

Herpesvirus transcription therefore occurs as a 3 stage cascade. A small number of α genes are expressed via the host transcriptional machinery, and the products of these genes are principally involved in starting the subsequent rounds of transcription. Following the synthesis of α proteins, β and γ1 proteins are then produced. Following synthesis of β proteins, viral DNA replication commences, which triggers the production of γ2 proteins. γ1 protein levels only reach their peak following the onset of viral DNA replication.

1.2.3.c Assembly and Release of Progeny Virus

Once viral DNA has been replicated and the structural proteins synthesised, they are incorporated into new virus particles inside the nucleus (reviewed in Roizman, 2001a). New viral genomes are packaged into preformed capsids, the exact

mechanisms for which are not fully elucidated, but this process starts with the procapsid structure (Sheaffer, 2001).

There follows a complex process of envelopment that appears at first glance to be rather inefficient (Whiteley, 1999; Skepper, 2001; Mettenleiter, 2002). Mature capsids bud through the inner nuclear membrane, which contains viral glycoproteins. These primarily enveloped capsids then bud through the outer nuclear membrane where the primary envelope is lost. The cytoplasmic capsids then associate with the various tegument proteins, including VP16 and vhs. These 2 proteins appear to interact with the capsids and help the final envelopment process. Final envelopment involves the mature capsids, with their associated tegument proteins, budding into exocytotic vesicles, which contain all the glycoproteins associated with the mature virions on their membranes. These now infectious virions either remain cell associated within these vesicles and spread to uninfected cells via virus-induced fusion, or are released from the cell in the exocytotic vesicles. It seems unlikely that any stage of the herpesvirus life cycle would have evolved to be inefficient, and therefore further study to examine the advantages of this mechanism is required.

There is one main alternative hypothesis for the stages following the primary envelopment step. This suggests that the primary enveloped capsids exit the cell via the Golgi apparatus and the associated secretory pathway without intermediate de-envelopment/re-envelopment steps (Johnson, 1982; Campadelli-Fiume, 1991). It remains to be determined conclusively, which pathway is correct.

1.2.3.d Latency

While persistence is a relatively common phenomenon of DNA viruses, latency is the hallmark of herpesviruses. The virus attaches and enters the host cell, but does not initiate the gene expression cascade characteristic of lytic infection. The viral genome is maintained in a genetically identical, but structurally different form, compared to that during lytic infection. Latent genomes form closed circles, usually as host-histone associated supercoiled episomes (Efstathiou, 1986; Deshmane, 1989). A very restricted programme of transcription is also usually observed, often localised to a small region of the genome (Stevens, 1987; Sugden, 1994).

The physiology and molecular biology of latency is one of the most varied aspects of herpesvirus biology. Each herpesvirus has evolved its own unique ecological niche within its host, and any host species can harbour multiple herpesvirus species. The alphaherpesviruses become latent in ganglionic neurones (Preston, 2000; Kennedy, 2002). HCMV establishes latency in bone marrow-derived myeloid

progenitor cells (Jarvis, 2002; Sissons, 2002). EBV resides latent in resting B cells (Crawford, 2001). The reservoir for latent HHV-8 is unclear, but based on its association with B cell and endothelial tumours, candidates include endothelial precursors (KS), pre- (multicentric Castlemans disease) and post-germinal centre B cells (primary effusion lymphoma) (Ambroziak, 1995; Cesarman, 1995; Du, 2001).

There are 3 stages to latency: establishment, maintenance and reactivation. Establishment involves infection of the appropriate cell type, followed by what appears to be failure of the productive gene cascade to be initiated. This is likely to be a function of the specific cell type in which latency occurs (Wagner, 1997). HSV-1 can maintain latency passively as it does not need to express any of its genes. It does, however, express a single latency associated transcript (LAT) for which no function, or even protein, has been elucidated (Stevens, 1987). The primary 8.3kb LAT is, however, spliced to give an abundant 2.0kb LAT, which is then further spliced to a 1.5kb LAT (Spivack, 1987; Garber, 1997). There is also a recent report showing that peptides, artificially synthesised using the largest open reading frame (ORF) of the LAT as template, become localised to distinct punctuate structures in infected cell nuclei, and can function to silence exogenous promoters inserted into the viral genome (Thomas, 2002). Therefore, although no LAT protein has been found, it certainly has potential to function in latency.

EBV is more active transcriptionally while in its maintenance phase of latency, perhaps related to its site of latency being a dividing cell as opposed to a post-mitotic neurone. EBV expresses a number of genes involved directly in maintaining the viral genome, as well as those involved in immune evasion and prevention of apoptosis (Sugden, 1994). In fact there are 3 classes of EBV latency, each with a different transcriptional profile, and related to the cells in which they are latent.

Reactivation of herpesviruses involves a switch to the productive transcription cascade, and is the least understood phase of latency. In particular, the molecular trigger for reactivation has eluded discovery. Studies have shown that HSV-1 reactivation efficiency is reduced if LAT is not expressed, but as no structural RNA or LAT protein has been observed, the mechanism behind this reduced reactivation requires further study (Bloom, 1994). HCMV appears to have a passive mechanism for reactivation. When the myeloid lineage progenitor cells, in which the virus is latent, differentiate into macrophages, the change in cellular environment results in conditions permissive for IE gene expression. The lytic gene expression cascade ensues and progeny virus is released (Sissons, 2002). However, the mechanisms of HCMV reactivation are still far from clear. For example it is unknown which

cytokines and pathways are responsible for the phenotypic changes that trigger reactivation.

1.2.4 Genomic Structure and Organisation of MHV-68

1.2.4.a Herpesvirus Genomic Structure

The HSV-1 genome is 152,260 base pairs (bp) in length and 68.3% G + C (McGeoch, 1988). The 2 covalently linked components of the genome are known as the long and the short. Both of these components contain unique sequences (U_L and U_S) flanked by inverted repeats (ab & a'b' around U_L, and ac & a'c' around U_S). The 2 components can invert relative to each other, which results in any population of HSV-1 consisting of 4 different genomic arrangements (Sheldrick, 1975; Roizman, 2001b). HSV-1's genome encodes at least 84 proteins, of which 14 are in U_S, 1 is in the repeat sequences around U_S, and 4 are in the repeat sequences around U_L, leaving 65 in U_L. Of these 84 genes, 39 have been characterised as essential for HSV-1 to complete its productive life cycle *in vitro* (Roizman, 2001a).

The betaherpesvirus subfamily includes the largest sequenced herpesviruses with many genomes of more than 200kb - chimpanzee CMV's genome is 241,087bp (Davison, 2003). Members of the *Roseolovirus* genus have smaller genomes, HHV-6's genome is 159,321bp (Gompels, 1995) and has 2 variants, HHV-6A and HHV-6B, which differ in the genetic and biological properties (Isegawa, 1999).

Gammaherpesviruses possess genomes in the same size range as HSV-1. EBV's genome is 172kb with 84 predicted s (Baer, 1984), while HHV-8's is 165kb with 81 ORFs predicted (Russo, 1996) and MHV-68's genome is 119kb with 80 predicted ORFs. Gammaherpesviruses have collinear genomes, possessing conserved blocks of genes, even between γ1 and γ2 herpesviruses. The majority of MHV-68's genes have homologues in HHV-8 and EBV (Baer, 1984; Russo, 1996; Virgin, 1997). The overall genomic organisation of 4 different gammaherpesviruses is shown in Figure 1.4. Virus-specific genes tend to be found interdispersed around regions conserved by all gammaherpesviruses (Chee, 1990). For MHV-68, the virus-specific genes to the left of ORF 4 and to the right of ORF 69 are of particular interest as they are homologous to host genes. While some of HSV-1's genome has been characterised functionally, the genes of gammaherpesviruses remain mostly uncharacterised. Instead genes have been assigned putative functions based on homology to characterised genes.

1.2.4.b MHV-68 Genomic Structure

The MHV-68 genome is fully sequenced and has been shown to consist of a 118,237bp unique region flanked by multiple copies of a 1,213bp terminal repeat sequence (Virgin, 1997). It is predicted to encode 80 gene products, 63 of which are homologous and collinear to their counterparts in Herpesvirus saimiri (HVS) and HHV-8, and designated ORF 4-69. These genes make up the conserved gene blocks that are characteristic of gammaherpesviruses as shown in Figure 1.4.

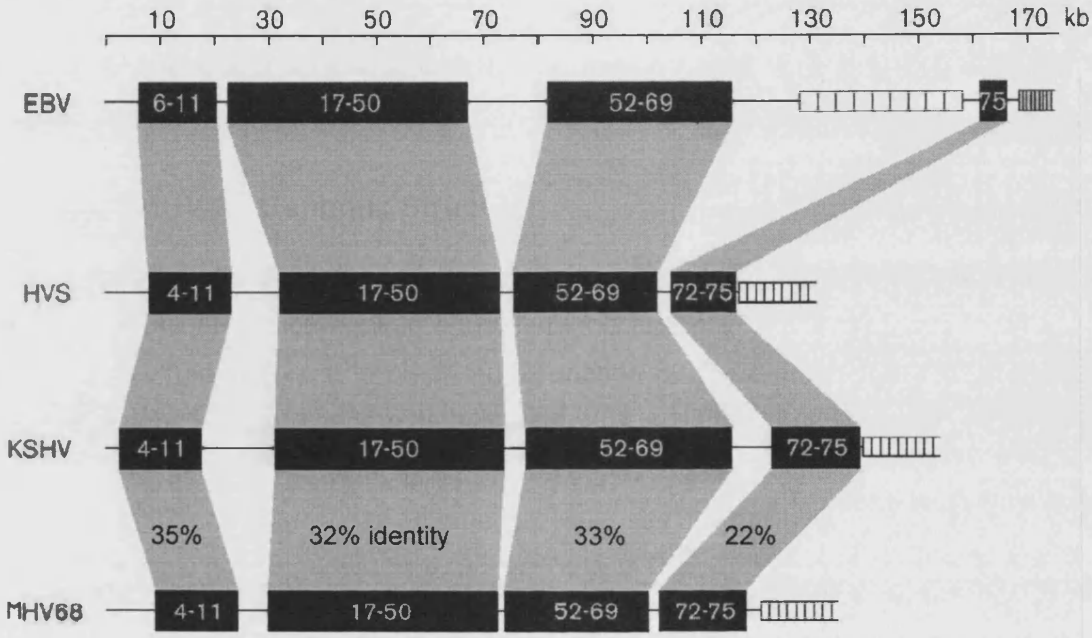


Figure 1.4 (preceding page). Comparison of gammaherpesvirus genome organisations. Blue blocks represent groups of genes, their constituent ORFs are shown in each block. Empty blue rectangles represent repeat regions. The name of each herpesvirus is given on the left in red: EBV – Epstein-Barr virus, HVS – Herpesvirus saimiri, KSHV – Kaposi's Sarcoma-associated herpesvirus or human herpesvirus 8, MHV-68 – Murine gammaherpesvirus-68 (adapted from Virgin, 1997).

1.2.4.b MHV-68 Genomic Structure

The MHV-68 genome is fully sequenced and has been shown to consist of a 118,237bp unique region flanked by multiple copies of a 1,213bp terminal repeat sequence (Virgin, 1997). It is predicted to encode 80 gene products, 63 of which are homologous and collinear to their counterparts in Herpesvirus saimiri (HVS) and HHV-8, and designated ORF 4-69. These genes make up the conserved gene blocks that are characteristic of gammaherpesviruses as shown in Figure 1.4.

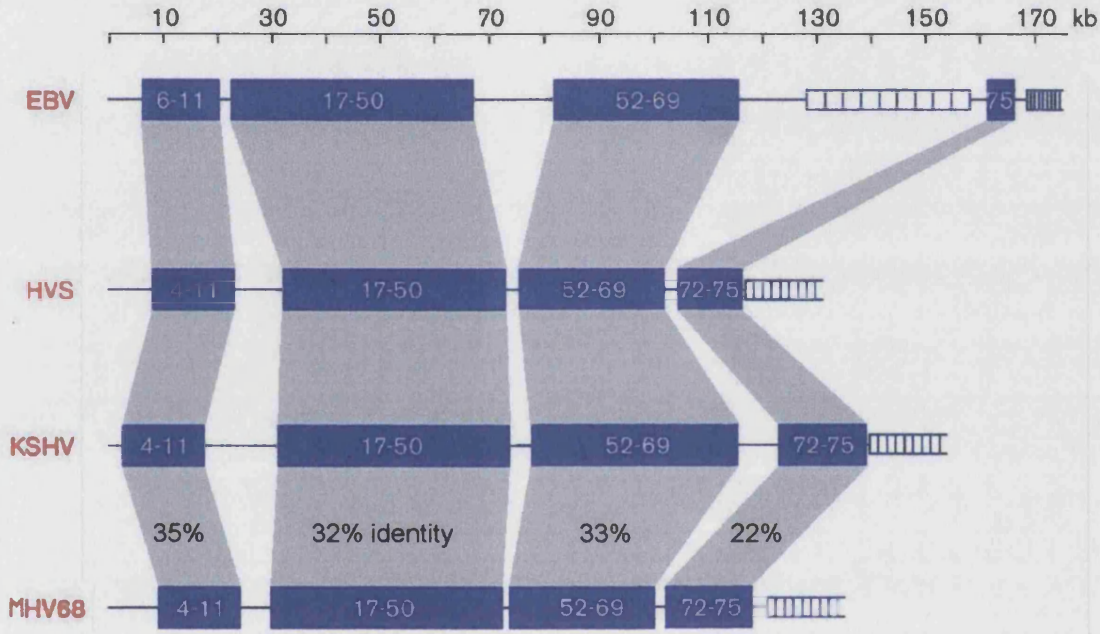


Figure 1.4 (preceding page). Comparison of gammaherpesvirus genome organisations. Blue blocks represent groups of genes, their constituent ORFs are shown in each block. Empty blue rectangles represent repeat regions. The name of each herpesvirus is given on the left in red: EBV – Epstein-Barr virus, HVS – Herpesvirus saimiri, KSHV – Kaposi's Sarcoma-associated herpesvirus or human herpesvirus 8, MHV-68 – Murine gammaherpesvirus-68 (adapted from Virgin, 1997).

Interdispersed among these homologous genes, lie MHV-68 specific genes designated ORF M1-M14 (the predicted ORFs of MHV-68 are listed in Appendix III). ORF M11, however, encodes a bcl-2 homologue common to gammaherpesviruses. It has been designated MHV-68 specific gene M11 as there is only a weak homology to counterparts in other gammaherpesviruses. M11 is found between ORF 72 and 73, at the right-hand end of the genome. This is the same position as the EBV bcl-2 homologue, but a different one to HHV-8's bcl-2 homologue. The gene in HHV-8 is named ORF 16 due to its position in the genome. It seems that in ancestral gammaherpesviruses, the bcl-2 homologue was at the right-hand end of the genome, and that some form of recombination event occurred in the HHV-8 lineage.

The genomic organisation of MHV-68 is shown in Figure 1.5. Generally speaking, herpesviruses are promoter-rich and so polycistronic transcripts are rare (Whitley, 2001). However, transcripts do often share polyadenylation sites, resulting in families of co-terminal transcripts. These sets of co-terminal transcripts also occur in the MHV-68 genome, for example ORF 4 and 6 share a polyadenylation site although they each have their own promoter (Kapadia, 1999).

As well as the terminal repeat, there are 2 internal repeat sequences. One is predicted to lie between positions 26778 and 28191 on the genome and consists of a 40bp repeated sequence. A second repeated sequence of 100bp lies between positions 98981 to 101170 (Virgin, 1997).

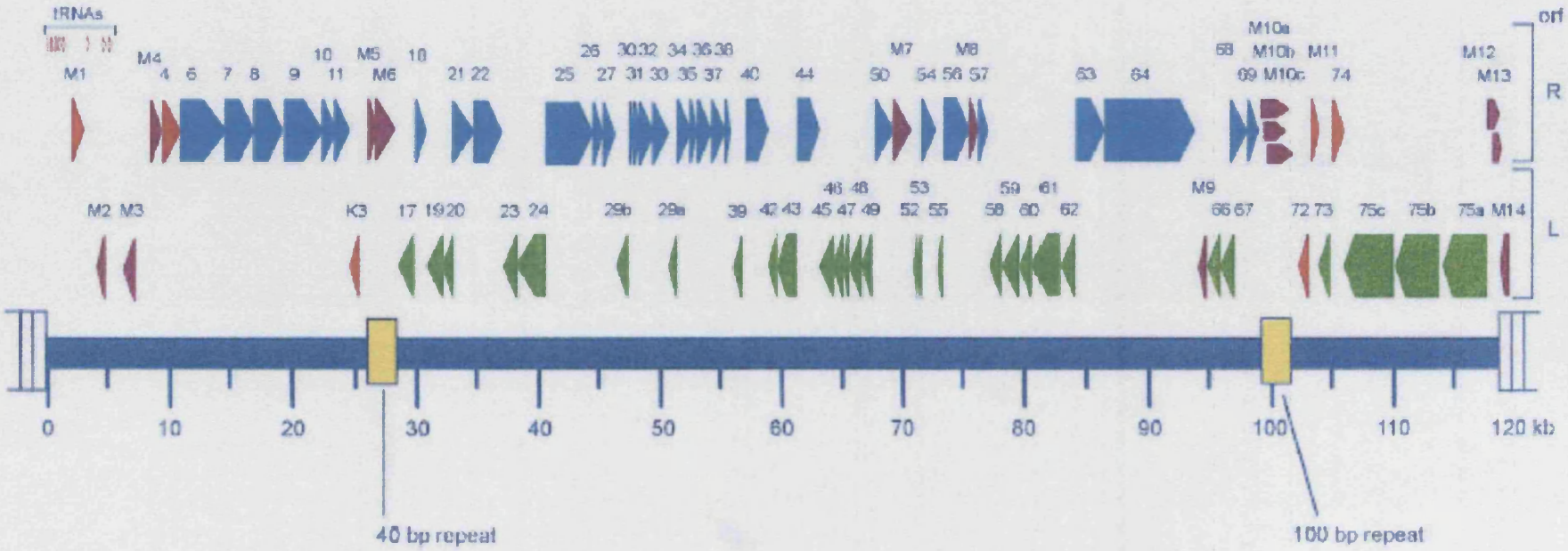


Figure 1.5. Genomic organisation and ORFs of MHV-68.

The MHV-68 genome's unique region is shown as a thick blue line. Terminal repeat units are depicted as open rectangles at the ends of the genome. Internal repeats are shown in yellow. ORFs homologous to genes found in HVS and HHV-8 genomes are depicted in blue or green depending their orientation. ORFs which are homologues of known cellular genes or of genes in non-herpesviruses are depicted in red. ORFs with no significant homology to any known cellular or viral gene product are depicted in purple. Also shown clustered near the left end of the unique region are 8 potential tRNA-like genes (adapted from Virgin, 1997).

At the left-hand end of the genome lie 8 novel sequences that encode tRNA-like sequences (Bowden, 1997). EBV and HVS are known to also encode families of non-coding RNAs with features typical of tRNAs (Lee, 1988; Swaminathan, 1991). These occur at similar loci to those of MHV-68, as all 3 lie close to their respective origins of replication. The function of these is unclear but in MHV-68, they are transcribed abundantly during productive and latent infection (Bowden, 1997). However, deleting 4 of the 8 tRNA-like sequences has no effect on replication *in vitro* or on establishment and reactivation of latency *in vivo* (Simas, 1998a). Similar observations were made for EBV and HVS (Murthy, 1989; Swaminathan, 1992). Based on these characteristics however, the MHV-68 tRNAs have been used successfully as markers of latency (Simas, 1999).

All gene names used in this thesis are based on the sequence analysis performed at the time that MHV-68 genome was fully sequenced by Virgin *et al.* (1997). However, MHV-68 has also been sequenced and analysed a second time (Milligan, 1998). The second analysis of the MHV-68 genome was more sophisticated than that performed in the original study, as Milligan *et al.* (1998) mapped features such as polyadenylation signals and potential splice sites.

In this analysis, several of the ORFs designated by Virgin *et al.* were demoted to the status of being unlikely to encode proteins. They are ORF M5, M6, M10 (a, b & c), M12, M13 and M14. There is little experimental data concerning any of these predicted ORFs. Furthermore, M8 was predicted to be an exon of ORF 57. Similarly ORF 29a was incorporated into ORF 29. ORF M9 was found to be homologous to a gammaherpesvirus capsid protein and has therefore been renamed ORF 65. For comparable reasons, M7 and K3 were renamed ORF 51 and ORF 12, respectively. Finally, three sequences were elevated to ORF status, ORF 17.5, 28 & 67a. By way of summary, a list of the ORFs of MHV-68 and their predicted functions can be found in 9.3 Appendix III. Furthermore any genes that were characterised at the time of the start of this thesis are also listed.

1.2.4.c MHV-68 Genes with Homologues in Other Viruses

The majority of MHV-68 genes are homologous to other gammaherpesvirus genes. Many of these genes have been assigned a putative function based on homology to characterised herpesvirus genes. ORF 21 encodes MHV-68's thymidine kinase (tk) which is non-essential *in vitro* as mutants of MHV-68 lacking it show no phenotype (Coleman, 2003). However, *in vivo*, titres in the lung are severely reduced by a factor of 1,000, so that almost no virus can be isolated. However, even with this severe restriction of productive viral replication, MHV-68 is able to infect

splenocytes and establish latency. Infectious centre assays show that levels of latent virus in the spleen are reduced compared to wild type (wt) initially, but by 1 month pi, levels of latency for the tk deficient virus reach that of wt virus.

It seems that deoxythymidine monophosphate (dTMP) levels in tissue culture cells are high enough not to hinder viral replication. While *in vivo*, the cells infected by MHV-68 in the lung do not appear to contain large reserves of dTMP. However, as DNA replication is not a feature of the establishment and maintenance of latency, levels of latent virus eventually reach the levels observed during wt infection.

There are also 2 genes that are homologues to non-herpesvirus genes. The protein predicted to be encoded by ORF M1 is homologous to the serpin proteins of the poxviruses, and unique to MHV-68 among the gammaherpesviruses. These proteins have been shown to be involved in processes such as control of host inflammatory responses and regulating apoptosis (Chua, 1990; Brooks, 1995).

Intraperitoneal (i.p.) infection of mice with insertion or deletion mutants with respect to M1, result in decreased splenic titres 9d pi, but levels of latent virus remain similar to that observed for wt infections. Interestingly the rate of reactivation was increased approximately 5-fold (Clambey, 2000). However, intranasal (i.n.) infection of mice with a deletion mutant of MHV-68 missing M1 and tRNAs 1-4 showed no phenotype both *in vitro* and *in vivo* (Simas, 1998a). Further analysis is required to elucidate whether the route of infection or the tRNAs contributed to the discrepancy in results.

ORF K3 is homologous to K3 of HHV-8 and the major immediate early protein (IE1) of BHV-4 (van Santen, 1991). The main point of homology between these proteins is a single motif that is completely conserved (Virgin, 1997). ORF K3 is transcribed both during productive infection, and during establishment of latency. Though not essential for the virus, it modulates the host immune response by down regulating MHC class I expression and helping evade the cytotoxic T lymphocyte response (Stevenson, 2002).

1.2.4.d MHV-68 Genes with Homologues in the Host Genome

Four cellular homologues were identified within the MHV-68 genome. ORF 4 is a complement regulatory protein homologue, possessing 4 regions with homology to many known complement regulatory proteins. This homologue is also found in HVS and HHV-8 (Albrecht, 1992; Russo, 1996). In particular, ORF 4 was found to have homology to murine decay-accelerating factor and human membrane co-factor protein. These proteins regulate serum complement via C3 convertase, as does

ORF 4 of HVS (Fodor, 1995) and HHV-8 (Spiller, 2003). ORF 4 of MHV-68 was found in the supernatant of infected cells, and on their cell surfaces, and has also been shown to inhibit complement activation (Kapadia, 1999). Both MHV-68 and HHV-8 encode 3 isoforms of the ORF 4 protein, at least one of which is soluble. A deletion mutant of MHV-68 lacking ORF 4 replicates *in vitro*, but with reduced growth characteristics (Adler, 2000), which suggests that the protein may have an additional role to that of complement regulation.

ORF 72 is predicted to encode a D-type cyclin. HHV-8, HVS and EBV also encode cyclin homologues, and EBV's homologue has been shown to regulate cellular cyclin D expression (Nicholas, 1992; Sinclair, 1995; Chang, 1996). This level of conservation through the subfamily indicates a conserved role for this gene. ORF 72 of HVS has been observed to be involved in binding cyclin-dependent kinase 6 (cdk6) (Jung, 1994), and it's counterpart in HHV-8, with cdk6 and weakly with cdk4 (Li, 1997). Although the mechanisms are unclear for MHV-68's cyclin D homologue, it has been classified as an oncogene due to the appearance of lymphoid tumours in transgenic mice expressing the gene (van Dyk, 1999). The authors further showed that it promotes cell cycling of thymocytes and inhibits their differentiation. Furthermore the gene appears to be involved in reactivation, as ORF 72 mutants of MHV-68 show reduced levels of reactivation (Hoge, 2000).

ORF M11 has weak homology to the bcl-2 homologues of EBV and HHV-8, which are known to inhibit apoptosis (Henderson, 1993; Chang, 1997). ORF M11 has been shown to possess antiapoptotic activity as well, however, it differs from the products of other gammaherpesviruses as it is digested by caspases (Bellows, 2000). The products of M11 caspase digestion, unlike the bcl-2 caspase digestion products, lack proapoptotic activity. The bcl-2 homologues of EBV, HHV-8 and bovine herpesvirus 4 (BHV-4) are not susceptible to caspase digestion and therefore retain their antiapoptotic properties by this strategy.

Finally, MHV-68 ORF 74 is homologous to the G-protein-coupled receptor family, in particular, the IL-8 receptor. HVS ORF 74 has been shown to function as an IL-8 receptor (Ahuja, 1993). HHV-8 ORF 74 has been shown to be a constitutively active G-protein-coupled receptor. It binds both α and β chemokines, and is capable of producing signals for cell growth (Arvanitakis, 1997). HHV-8 ORF 74 is also found to be expressed in tumours cells (Cesarman, 1996).

Recombinant MHV-68 lacking ORF 74 shows no phenotype *in vitro*. However, the reactivation efficiency of this virus, from latently infected splenocytes, is reduced.

Also while CXC chemokines increase wt MHV-68 replication, the ORF 74-deficient recombinant does not respond to these chemokines (Lee, 2003).

1.2.4.e MHV-68 Unique Genes

MHV-68 encodes a number of genes that have no known homologues, designated ORFs M1-M14 (Virgin, 1997). Analysis of the predicted amino acid sequence for ORF M2 returned no homologues. Studies of gene expression in the S11 B cell tumour cell line prompted Husain *et al.* (1999) to suggest that M2 was transcribed during latency.

The S11 line was originally derived from lymphomas of MHV-68-infected mice, and has been shown to contain episomal forms of MHV-68's genome at a ratio of 1:3.0 to the genome in it's linear form (Usherwood, 1996a). As the linear form of the genome is associated with productive infection, lytic gene expression would also be expected to be occurring in this cell line. Indeed infective centres assays showed 2-4% of cells gave rise to MHV-68 antigen positive plaques, after 12 weeks in culture. This rose to 20-30% by 9 months in culture.

Husain *et al.* (1999) detected transcription of ORF M2 in the S11 line by northern blot analysis. It was the only abundantly expressed viral gene that was observed in the S11 cell line. However, when the S11 cell line was chemically induced to reactivate, levels of M2 transcripts were seen to rise. ORF M2 transcripts were not detected though in lytically infected BHK cells.

RNA protection assays of MHV-68 gene expression show that the M2 transcript is expressed in murine 3T3 fibroblasts during lytic infection. Furthermore M2 expression was not detected in latently infected splenocytes by this method (Rochford, 2001). The picture is further obscured by observations made following i.p. infection of mice, as a latent infection of peritoneal macrophages is observed with this routes of infection. These latently infected macrophages also express M2 (Virgin, 1999). However, it is not clear how significant these findings are as the i.p. route of infection is not a natural one.

Studies using mutant MHV-68 strains with disruptions in M2 have shown that growth curves *in vitro* are identical to those of wt virus. Similarly *in* vivo, the viral titres recovered from lungs of mice infected with the mutant strain are equivalent to wt infection titres. In the spleen however, *ex vivo* reactivation assays show that infections with the mutant strain show reduced levels of virus recovered. Therefore, it appears that M2 does play a role in latency, or at least reactivation from latency (Jacoby, 2002). Interestingly, if mice were infected i.p. as opposed to i.n., the levels

of reactivating virus from the spleen were equivalent for wt and mutant strain virus infections. This suggests that injecting virus directly into the peritoneum may compensate for M2 function. Further studies are now required to elucidate the role of ORF M2 during the virus' lytic and latency phases, and also the switch between them.

Transcription of ORF M3 produces an unspliced 1287bp polyadenylated mRNA that encodes an abundantly secreted protein (van Berkel, 1999). The M3 protein binds a broad range of chemokines including CC, CXC, C and CX_3C, although it shares no homology to known chemokine receptors (Parry, 2000). Transgenic mice expressing M3 and the chemokine CCL21 in the pancreas have been used to study the effects of M3 *in vivo*. CCL21 expression by the pancreas would normally result in development of lymphoid aggregates that resemble lymph nodes. While 18% of mice expressing just CCL21 did show lymphoid aggregates, only 2% of the mice expressing both M3 and CCL21 showed similar symptoms (Jensen, 2003). This study demonstrated that M3 can block the activity of a CC chemokine *in vivo* and suggests that M3 could disrupt the mobilisation of lymphoid cells regulated by CCL21 as an immune evasion strategy.

ORF M7 encodes MHV-68 glycoprotein 150 (gp150), a protein that is well conserved through the herpesvirus family. It has be shown to be present on the surface of infected cells, and in virus particles (Stewart, 1996). A recombinant vaccinia virus expressing MHV-68 gp150 has been constructed and used to study potential vaccines against MHV-68 infection (Stewart, 1999).

1.2.5 Epidemiology and Pathogenesis of MHV-68

Seropositivity of wild rodents to the murine herpesvirus in Slovakia was found to vary from 1-12.5% depending on the region (Mistrikova, 1985). In the north of England, seropositivity was 13% for the wood mouse (Blasdell, 2003). Although murine herpesvirus isolate 68 had originally been isolated from the bank vole, that species was only 2.7% seropositive in the north of England. This suggests that the murine herpesvirus can infect wood mice and bank voles with similar efficiencies. It would be interesting to test for seropositivity in other rodents.

Following i.n. infection, the lung is known to be the primary site of infection (Sunil-Chandra, 1992a). Infected mice tend to develop largely subclinical infections, although there is a marked splenomegaly associated with a characteristic increase in the number of germinal centres (Sunil-Chandra, 1992a). It seems likely that naturally occurring MHV-68 infections are generally sub-clinical.

The productive infection in the lung peaks around 6d post-infection (pi; Sunil-Chandra, 1992a). The virus is found associated with alveolar epithelial and mononuclear cells in the lung. Infection then resolves 10-12d pi. Virus can sometimes be isolated from the various other organs during this phase of infection, but generally at lower levels (Svobodova, 1982a).

The splenomegaly and an associated swelling of lymph nodes marks the establishment of a population of latently infected B cells, which peaks around 2-3 weeks pi (Sunil-Chandra, 1992b; Bowden, 1997). The number of infected lymphocytes eventually falls to around 1 latently infected cell per 10^{5-6} B cells (Nash, 1994). Interestingly, there is evidence that MHV-68 establishes a persistent infection, rather than a truly latent infection (Christensen, 1999). Murine myeloma B cells infected with MHV-68 show no cytopathic effect (cpe) but produce infectious virus. This can be inhibited with acyclovir, but virus remains in the cells as determined by infectious centre assay (Sunil-Chandra, 1993). Furthermore, while the CD4+ T cell population peaks at 1:50 during the acute response and then falls, they remain activated and at stable numbers (1:400 – 1:500) for 14 months pi, presumably by continuously reactivating, persistent virus (Christensen, 1999).

It has been suggested that the lungs are a secondary site of latency, as viral DNA has been detected in the lungs of infected mice by PCR, 12 months pi (Stewart, 1998). Furthermore, both linear and episomal forms of the viral genome were also detected in lung tissue. However, the same study also detect viral DNA occasionally in the blood and bone marrow of mice. The source of viral DNA could therefore be circulating B cells that are latently infected. If this is the case then it seems likely that the source of viral DNA in the lungs could also be latently infected B cells. The same group also reported that viral DNA was detectable by PCR in lung tissue of μMT mice 35d pi (Usherwood, 1996b). As these mice do not have B cells, they could not be acting as template for the PCR. These results suggest that the lungs are a second site of latency. However, the progress of MHV-68 infection in immunocompromised μMT mice may not be the same as seen in a natural infection. Therefore, further analysis of the latent phase of MHV-68's life cycle is required before it can be fully elucidated.

The human gammaherpesviruses are associated with a number of lymphoproliferative diseases. Long term persistence of MHV-68 has also been associated with lymphoproliferative disease, with one study showing 9% of animals developing lymphomas over a 3 year period, half of which were high-grade (Sunil-Chandra, 1994). All the tumours were a mixture of B and T cells, however, and not

every malignant B cell was MHV-68 DNA positive. In fact, positive cells were infrequent, reminiscent of EBV and HHV-8

1.3 Studies of Gene Expression in the Post-Genomic Era

1.3.1 Transcriptional Profiling

The regulated expression of genes directs all biological processes including many of the pathological changes that occur during disease (Muralidhar, 1998; Dietz, 2002; Carter, 2002; Jones, 2003; Levine, 2003). Analysis of gene expression can therefore elucidate one of the first steps towards producing a phenotype. This analysis can be aided by prior knowledge of the transcript population, which can be predicted from genomic sequence data (Russo, 1996; Virgin, 1997). There has been a recent wealth of genomic sequence data, driven partly by the high profile human genome sequencing project (Venter, 2001). Global analyses of transcription have been developed to take advantage of this data, and are particular in that they are designed to examine all gene expression occurring in a given sample, as opposed to focusing on a single or few genes (Spellman, 1998).

1.3.2 Differential Display

The differential display system, developed by Liang and Pardee (Liang, 1992; Liang, 1993), is often used in preference to the complementary DNA (cDNA) subtraction library technique. cDNA subtraction involves hybridising RNA from one sample against cDNA from another sample. The majority of RNA species will hybridised to complementary cDNA sequences. Any single stranded sequences that remain will represent genes that are differentially expressed in the 2 samples. The differential display system has a number of advantages over the subtractive hybridisation methods as it is technically simple, yet more sensitive, faster and more reproducible. Also unlike other techniques whose success is biased by relative mRNA abundances, differential display detects both abundant and rare mRNAs, on the condition that there is an arbitrary primer sequence that matches the target mRNA.

A further advantage of differential display is that prior knowledge of the geneset of an organism is not required. Instead, it simply isolates transcripts that are differentially expressed in one sample population, relative to another. The system has been used successfully to identify and study the effects on gene expression during such diverse events such as: transformation of fibroblasts with the Tax protein of human T cell leukaemia virus (Shimizu, 2003), infection of epithelial cells

with pneumovirus (Domachowske, 2001) and the effects of nitric oxide on the apoptotic pathways of hepatocytes (Zamora, 2001).

It has also been used to reveal the cellular changes that occur when latent HSV-1 is stimulated to reactivate (Tal-Singer, 1998). In this study, RNA was extracted from the trigeminal ganglia of latently infected and control mice, and differential display performed. As well as some novel sequences, a number of genes involved in the interferon pathway were up-regulated in samples with reactivating virus. Although a large proportion of the differential expression could be attributed to the stress of explantation, it seems likely that some of the pathways involved in this response would also be common to reactivation. The authors suggested that the interferon pathway would be a good target for further study of HSV-1 reactivation.

These results indicated how the cellular environment, from a transcriptional perspective, differed when HSV-1 reactivates. If a more natural model of reactivation could be used, the results would become more representative, as the transcriptional noise from the explantation stress could be reduced.

The lytic cycle of herpesviruses has also been studied. Primary human foreskin cells were infected with HCMV and RNA isolated 8h pi (Zhu, 1997). The differential display returned 57 partial cDNA clones, which corresponded to 26 differentially, expressed mRNAs. Eleven of these were virally encoded, while 15 were of cellular origin. Of these 15 up-regulated cellular genes, the 11 genes that could be identified and the remaining previously uncharacterised genes were all shown to be inducible by interferon α. However their up-regulation during infection occurred even when antibodies against interferon α and β were present in the medium.

The results suggested that a constituent of the viral particle could have been responsible for inducing the up-regulation of the 15 cellular genes. The differential display suggested that elements of the interferon response pathway were up-regulated by viral infection. Further study is now required to confirm whether expression of these genes were triggered by the virus, or by the cell in response to the virus.

An overview of the differential display protocol is shown in Figure 1.6. Having isolated RNA from the relevant samples, the RNA is reverse transcribed using anchored oligo-dT primers. There follows a PCR amplification for the 3' terminal regions of the poly-adenosine (poly-A) RNAs by using an anchored primer (with a 3' string of T residues) and one degenerate primer. Using enough degenerate primers in combination with the anchored ones will allow the entire poly-A RNA population to be amplified (Liang, 1993).

Differential gene expressions are visualised by separating the products of each amplification out on a large DNA sequencing gel. By comparing the different lane profiles of the two RNA samples, and in particular the presence of a band in one lane while being absent in another, inferences can be made concerning the differential expression of the gene represented by a band. For example, in terms of a viral infection, a band absent in an uninfected sample lane but present in the infected sample lane would be indicative of a gene that is being expressed during infection, and therefore most likely a viral gene. This differential expression can be confirmed by northern blotting and the gene represented by the band identified by sequencing.

Essential to this technique are the primers used for amplification. It is necessary that each primer pair will amplify a subset of the total cDNA population, which can then be resolved on a gel. Also, there must be sufficient combinations of primers to allow coverage of the whole cDNA population.

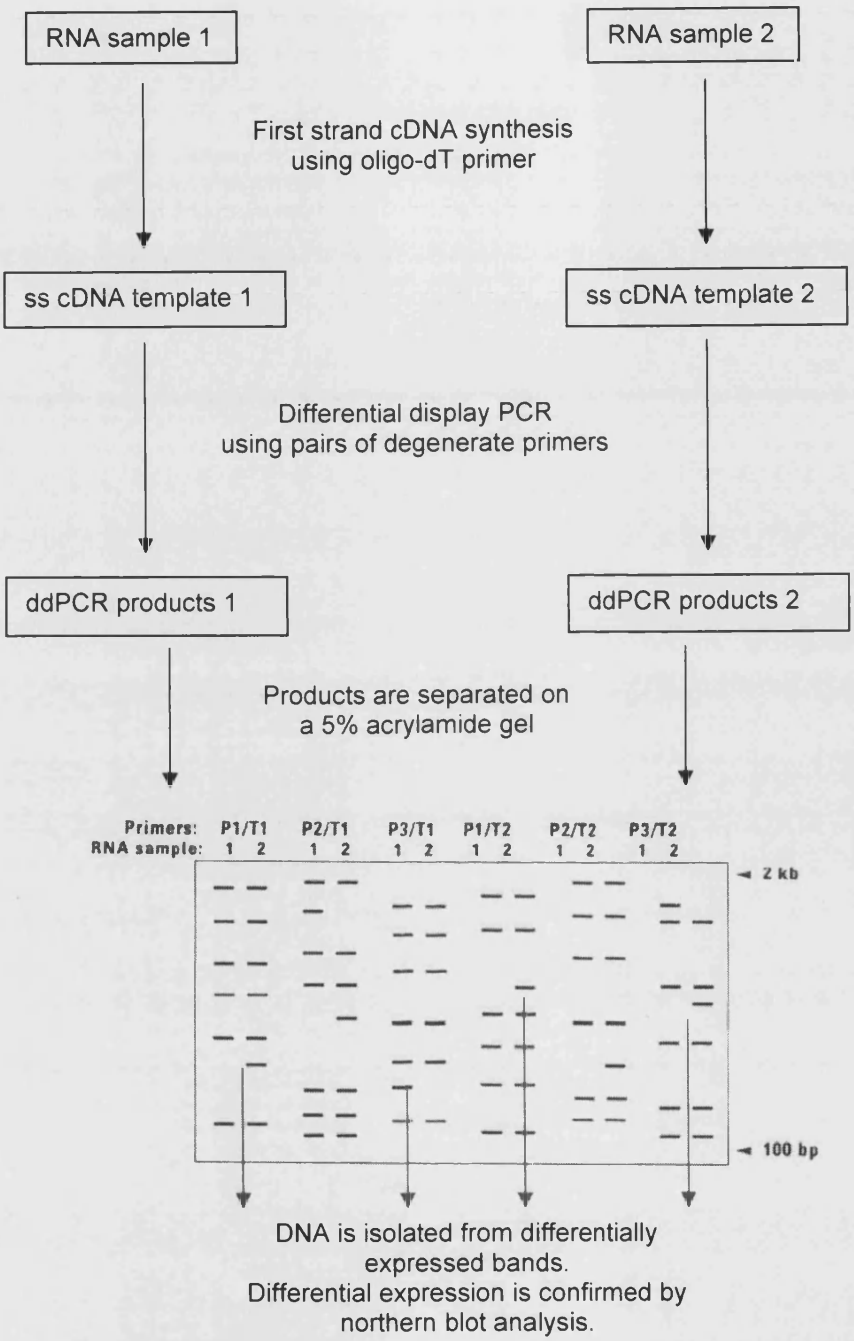


Figure 1.6. Overview of the differential display technique.

1.3.3 Array Technology

1.3.3.a Overview

The most widely implemented genomic technology at present is that of arrays (Schena, 1995). They are also referred to as microarrays, macroarrays, cDNA arrays, DNA arrays and oligo arrays, usually depending on the variant of array technology used. All arrays, are based around the same technology, that of a large collection of probes attached to a solid surface, used to detect transcript abundances within a sample (Southern, 1999). Arrays are currently the technique that most fully implements genome sequence data. Unlike the differential display system, arrays require knowledge of genome sequence and organisation, in order to design probes for each possible transcript. They are therefore dependant on accurate predictions of ORFs and transcript organisation to design probes. This approach has the advantage over differential display in that it is unnecessary to sequence and identify differentially expressed transcripts highlighted by the system. Each probe is tested using various algorithms at the design phase to reduce the chances of redundant probes, poor annealing chemistry and inter-probe interference (Mitsuhashi, 1994; Chan, 1995). In other words, more optimisation at design is possible than with most other technologies.

An overview of array technology is shown in Figure 1.7. Probes can be BAC clones containing appropriate inserts (Greshock, 2004), or cDNA fragments synthesised by PCR amplification (Duggan, 1999), or long oligonucleotides synthesised chemically (Lipshutz, 1999). The resulting arrays are known as DNA arrays and oligo arrays, respectively. The probes for the array are attached to a solid medium, usually a glass slide for oligo arrays or nylon membrane for DNA arrays. The resulting array of probes will bind complementary sequences from a pool of DNA.

Oligonucleotide probes tend to be smaller in size than cDNA probes and therefore can be arrayed at a higher density. cDNA probes are limited in that if PCR amplification of a DNA sequence is not possible, then it is difficult to produce a probe for the region of DNA. While it is possible to clone these fragments and then cut out the probe sequences from plasmids using restriction endonucleases, the yield of this technique is insufficient for the needs of an array. The cost of producing cDNA probes is far less than that of oligonucleotide probes, due to the expense associated with synthesis of long oligonucleotides. However, oligonucleotide probes that are synthesised using phosphoramidite chemistry (as opposed to the more common methods of soft lithography and optical deproteination; Lipshutz, 1999) have the advantage that the probeset can be altered simply by using a different

software template (Southern, 1999). Therefore, phosphoramidite chemistry-based probe synthesis provides flexibility not found in other systems, and is particularly useful for the optimisation of array systems. Phosphoramidite chemistry also allows the probes to be synthesised directly *in situ* on the glass slide, attached only at the end, which is likely to allow higher specificity than covalently bonding the entire probe to the glass slide (Southern, 1999; Cheung, 1999).

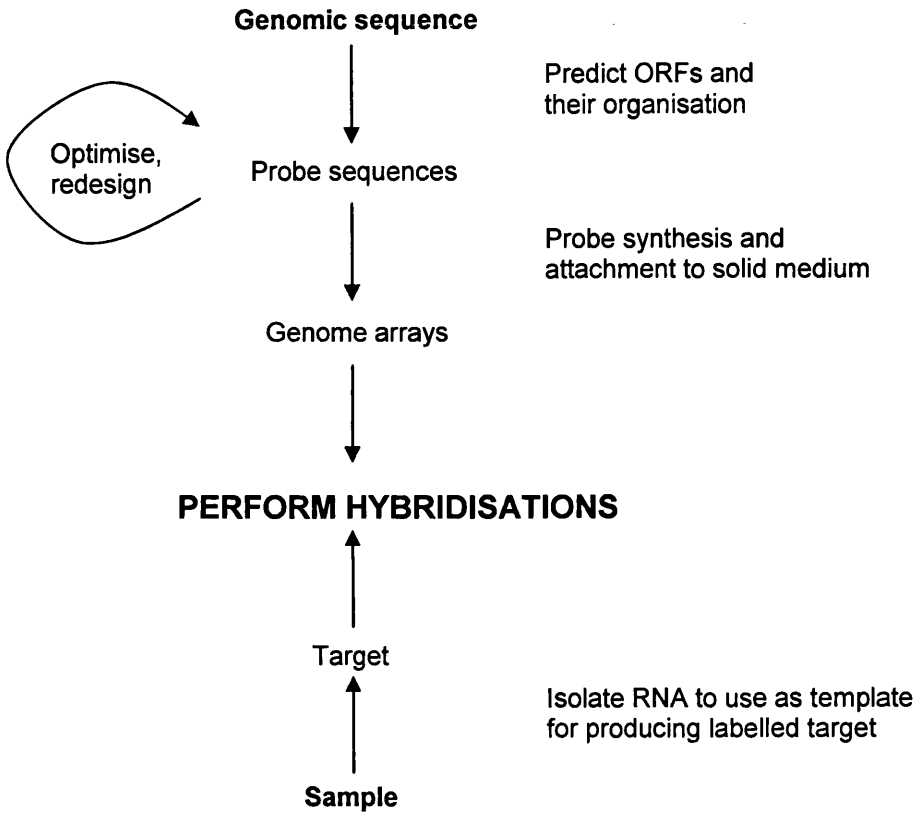


Figure 1.7. Overview of array design and experimental procedure.

Typically, mRNA is reverse transcribed using a labelled nucleotide to produce a labelled cDNA population representative of the original mRNA population. These target cDNAs are then hybridised to arrays, which are then washed stringently to remove any non-specifically bound cDNA. The hybridisation pattern can then be visualised and quantified in an appropriate manner depending on the label used.

If a radioactive label is used then samples are labelled and hybridisations performed singly. The results from each array are compared to those from other arrays. If fluorescent labels are used, it is also possible to label 2 target populations

with different colours and perform a dual colour hybridisation. This has the advantage of allowing a direct comparison between 2 samples hybridised to the same array, under identical experimental conditions. When comparing the hybridisation results of different arrays, as with radioactively labelled target arrays, strict controls must be used to take into account variation in experimental conditions, such as differences between arrays, differences in labelling and hybridisation efficiencies, and differences in radioisotope activities.

1.3.3.b Array Data Analysis

The results of arrays are quantified by a phosphoimager or laser scanner, depending on whether radiolabelled or fluorescently labelled nucleotides have been used. The amount of labelled cDNA target bound to each probe of the array is recorded as a signal intensity. A variety of control measurements, such as background and binding pattern of target to probe, may be taken as well to provide indications of the data's reliability.

As arrays tend to be designed to perform genome-wide analyses, the datasets produced tend to be large. One study attempting to identify prognostic markers for prostate cancer used an array of 9,984 probes (Dhanasekaran, 2001). Forty-four experimental and 2 control samples were analysed resulting in 459,264 individual data points. Viral analyses tend to produce smaller datasets due to the smaller size of their genomes, however even a simple study of HSV-1 transcription at 4 time points produced 400 data points, not including controls or repeats (Stingley, 2000).

Interpreting this magnitude of data has become a subject of research itself (Hughes, 2001). Although a number of computational analysis tools and packages have been introduced (Quackenbush, 2001), the standard method for analysing large sets of transcripts and experiments is clustering analysis (Eisen, 1998). This allows transcripts to be placed on a hierarchical tree showing the relatedness of transcription profiles to each other. Therefore, genes with similar expression patterns are clustered into groups, which branch back to form larger groups, much like phylogenic trees used for classification of species.

A key advantage of global gene expression studies is that as well as providing information about a single gene's expression, this transcription profile can also be analysed within the context of the global transcriptional landscape. Furthermore, global changes in transcription can also be analysed. HCMV is known to affect the transcription of many cellular genes, including interferon-induced genes, as shown by arrays (and also differential display; Zhu, 1997; Zhu, 1998). Array analyses have also shown that the global changes in cellular gene expression caused by HCMV

infection, are the same as that seen when the same cells are stimulated by type I interferons (Simmen, 2001). Indeed, the binding of HCMV glycoprotein B alone is sufficient to produce the same changes in cellular gene expression. Although the human gene array used in this study did not cover the entire human genome, 258 (of 6,600 on the array) genes were differentially expressed, and it therefore seems likely that the binding of glycoprotein B is a key part of inducing the cellular response to HCMV infection.

1.3.3.c Applications of and Conclusions from Array Studies

Transcriptional profiles of a number of herpesviruses have been published including those of Chambers *et al.* (1999), Stingley *et al.* (2000), Paulose-Murphy *et al.* (2001), Jenner *et al.* (2001) and Sciortino *et al.* (2001). There have also been a number of array-based studies investigating various conditions of infection. For example, the transcription of HSV-1's genes has been studied at 2h and 8h pi (Stingley, 2000). Two different global gene expression profiles were seen at these time points, with structural genes expressed predominantly at the later time point. Therefore, the virus regulated the expression of its genes, as has been shown previously (Jones, 1979). Categorisation of HSV-1 transcripts into immediate-early, early and late kinetic classes was attempted by using cycloheximide (CX) and phosphonoacetic acid to block protein synthesis and viral DNA replication, respectively (Stingley, 2000). While the results agreed with previous studies, this was interesting as it was the first time that this investigation had been performed at the individual transcript level across the whole genome, as opposed to at the protein level (Honess, 1974).

Arrays therefore allow a rapid global analysis of transcription and by using samples taken at various times pi, it is relatively simple to produce a temporal map of the transcription of each gene of the genome. By blocking *de novo* protein synthesis or viral DNA replication, it is also possible to simply categorise viral genes into their kinetic classes of expression.

There have been further analyses of immediate-early gene expression centred on HSV-1 VP16 (Yang, 2002). This protein is carried into the cell as part of HSV-1's tegument, and activates transcription of the immediate-early genes (reviewed in O'Hare, 1993). When a mutant strain of HSV-1, lacking VP16's activation domain, was used to infect cells, all transcript levels were lower than that seen in wt infections, and the overall pattern of transcription was aberrant. However, if cells were infected by the same mutant strain at higher multiplicities of infection (moi), the resulting transcription profiles were far closer to those of wt infections,

especially at the later time points. While it is a well noted phenomenon that infections using high moi's can sometimes overcome defects in viral gene expression, it was very interesting to see this as a function of the viral transcriptome, and in particular, the differences between wt and mutant strain infections.

Furthermore, while the expression profiles at the later time points were close to those seen with wt infections, at earlier time points, there were a number of significant differences. Yang *et al.* (2002) used the mutant strain lacking VP16 to model the cellular environment of latent infections of neurones by HSV-1, when tegument proteins are not present to initiate the transcription of immediate-early genes. They suggested that transcriptional profile observed when infecting cells with the VP16 mutant strain, would be similar to the transcription profile occurring during reactivation of HSV-1. In addition, they suggested that the low moi infection modelled a spontaneous reactivation and that the high moi infection modelled an induced reactivation. This shows the ability of arrays to model infections under a variety of conditions. It also illustrates the ability they give to dissect phenotypic events at the level of global gene expression.

VP16 is a transcriptional activator and primarily responsible for inducing viral gene expression during productive HSV-1 infection. Discerning the system-wide changes in transcription that occur when VP16 is not functional is of great help in elucidating the mechanisms of HSV-1 usurpation of the host cell. Arrays can also be used to characterise mutant strains in terms of changes in global gene expression. In this way they can simply assess the impact of any particular mutation on the transcription of other viral genes.

HHV-8 transcription has been studied and hierarchical clustering performed using a primary effusion lymphoma-derived cell line as a model system (Paulose-Murphy, 2001; Jenner, 2001). As this cell line is latently infected by HHV-8, latent gene expression was observed. However, the virus was also induced to reactivate, to allow analysis of lytic gene expression. Transcription profiles were clustered hierarchically, which sorted genes into functionally-related groups. Although few of HHV-8's genes have been elucidated functionally, putative functions have been assigned to many genes based on sequence homology. Cluster analysis of reactivation-associated expression profiles returned several functional groups, which confirmed the homology-based functional predictions. The array analysis was able to confirm these predictions in a simple and rapid manner. Furthermore, as

those genes with no putative function assigned were also clustered, predictions could be made as to the functional role of these wholly uncharacterised genes.

1.3.3.d Disadvantages of Array Analyses

Arrays can be thought of as performing a large number of simple DNA:DNA hybridisations simultaneously on a convenient single platform. They possess the weaknesses of any DNA hybridisation based technique (principally related to non-specific hybridisation), although by incorporating multiple probes it is possible to overcome some of the limitations and ambiguities that occur when using a single probe. One array used to analyse HSV-1 transcription consisted of 99 probes (Stingley, 2000), however due to the presence of overlapping transcripts, several of these probes bound more than a single transcript. This redundancy in probe specificity is unavoidable if transcripts overlap. However, probes can be designed to anneal to gene-specific regions of a transcript where possible.

Although there is no need to search the gene databases to identify results as with differential display, arrays are limited in a similar fashion in that if the sequence of the gene is not known, then it is not possible to design a probe to it. Therefore, many of the arrays used to study transcription consist of partial gene sets.

Arrays are not necessarily quantitative. Depending on the labelling method and probe design, they can be semi-quantitative. Proven quantitative techniques, such as northern blot analysis, are required for quantification of results. For example they can be used effectively to examine relative abundances of transcripts between samples. Distinct types of diffuse large B cell lymphomas have been successfully differentiated by their relative gene expression patterns (Alizadeh, 2000). However, absolute quantification of transcript abundances has not been a feature of array analyses in general. Instead array analyses tend to present their results as relative values, or induction/repression relative to steady state levels of gene expression.

Arrays have however been used successfully for prognostic purposes. Gene expression profiles of breast cancer have been used to produce a "poor prognosis" signature that can be used as an indicator of disease outcome and also to highlight patients that might benefit from more severe forms of therapy, such as adjuvant chemotherapy (van 't Veer, 2002). While accurate prognosis can help reduce the costs incurred by unnecessary treatments of patients, arrays are still an expensive technology and can be technically complicated, especially when interpreting results. Therefore, further development is required before they can be adopted by the health services.

1.4 Aims

The aim of this thesis is to further the characterisation of MHV-68 gene expression through the lytic phase of its life cycle. Initially, differential display and array systems will be developed to enable a global analysis of gene expression during MHV-68 infection. The array systems will be focused on viral transcription, while the differential display system will detect changes in both host and viral gene expression. A transcriptional study of MHV-68 will then be performed at the genomic level, to produce a series of expression profiles covering the MHV-68 genome.

MHV-68 is a recently characterised murine gammaherpesvirus. Its somewhat atypical properties, such as a rapid replication cycle and the ability to productively infect a wide range of cell lines, recommends its use as a practical small animal and *in vitro* system with which to model other, less amenable gammaherpesviruses. While the latent phase of gammaherpesvirus infections has been investigated with some success, productive lytic infections remain largely unfamiliar. Therefore, the events that occur during primary productive infection are of particular interest in this study.

The foundation of information that such a study can provide will prove to be a valuable tool in the further development and understanding of gammaherpesvirus biology.

2 Materials and Methods

2.1 Tissue Culture

2.1.1 General Conditions

All cell lines were maintained in humidified incubators, at 37°C with 5% CO_2 and grown as monolayers in 175cm^2 flasks (Falcon), unless otherwise stated. Cells were seeded at approximately 2 x 10^6 cells in 40ml of medium. If cells were cultivated in 6-well plates (Falcon), they were seeded at 8 x 10^5 cells per well. If cells were cultivated in roller bottles (Fisher), they were seeded at 2 x 10^7 cells per bottle. All media were supplemented with 100U/ml penicillin, 100µg/ml streptomycin and 2mM L-glutamine, unless otherwise stated. Growth media were further supplemented with 10% foetal bovine serum (FBS, European, Invitrogen) or newborn calf serum (NCS, European, Invitrogen) as indicated. Maintenance media were supplemented with 2% of the respective sera. All cells lines were mycoplasma negative.

2.1.2 NIH 3T3 Murine Embryonic Fibroblast Cell Line

The NIH 3T3 murine (*Mus musculus*) embryo fibroblast cell line (ATCC CCL1658) was supplied by the American Type Culture Collection in a frozen ampoule. NIH 3T3 cells were grown in Dulbecco's modified Eagle medium (DMEM, high glucose, without L-glutamine, Invitrogen) supplemented with 10% NCS. Cells were passaged at approximately 80% confluency to prevent the loss of contact-inhibited growth characteristics.

2.1.3 BHK-21 (Clone 13) Baby Hamster Kidney Fibroblast Cell Line

The baby hamster kidney fibroblast (BHK-21) cell line (Clone 13, ECACC 85011433; BHK cells) was supplied by the European Collection of Cell Cultures in a frozen ampoule. BHK cells were grown in Glasgow minimal essential medium (GMEM, with sodium bicarbonate, without L-glutamine; Sigma) supplemented with 10% FBS and 10% tryptose phosphate buffer (29.5g/l, Sigma). These were seeded at approximately 3 x 10^6 cells in 40ml.

2.1.4 Subcultivation of Cell Lines

Cell lines were subcultured by discarding the growth medium and rinsing the monolayer in PBS without magnesium or calcium (Invitrogen). Cells were dissociated by adding 5ml of 0.05% trypsin and 0.02% EDTA and incubating at

37°C in a humidified incubator at 5% CO_2 for about 1min, being careful to minimise the duration of digestion. Complete medium with serum was then added and a single cell suspension prepared by repeated pipetting. Live cell counts were determined using a haemocytometer and trypan blue staining.

2.1.5 Cryopreservation of Cell Lines

Cells were collected by centrifugation at 500g for 5min and resuspended at 5-10 x 10^6 cells/ml in standard growth medium, which was supplemented with 20% serum and 10% DMSO. Volumes of 1ml were aliquoted into cryopreservation vials and frozen at a cooling rate of –1°C/min using a Cryo freezing container (Nalgene) placed at –80°C for 16h. Frozen samples were then stored in liquid nitrogen.

2.1.6 Metabolic Inhibition of Cell Lines

2.1.6.a Cycloheximide

A stock solution of CX (Calbiochem) was made up in ethanol to a concentration of 10mg/ml and stored at –20°C, for up to 6 months. A concentration response curve was calculated for CX inhibition of protein synthesis in NIH 3T3 cells over 0.1ng/ml - 100µg/ml (2.1.7). A concentration of 2µg/ml CX was added to cell media, as this resulted in >95% inhibition of protein synthesis. Cells were pre-treated with CX for 30min prior to use in all experiments involving protein synthesis inhibition.

2.1.6.b 2'-deoxy-5-ethyl-β-4'-thiouridine

A stock solution of the antiviral thionucleoside analogue 2'-deoxy-5-ethyl-β-4'-thiouridine (4'-S-EtdU, a generous gift of the late Professor R. Walker) was dissolved in double distilled water (ddH_2O) to a concentration of 1mg/ml, at 60°C for 1h. The solution of 4'-S-EtdU was added to cell media to 200ng/ml (Barnes, 1999) and used to overlay cell monolayers to inhibit viral DNA replication. Cells were pre-treated with 4'-S-EtdU for 30min prior to use in all experiments involving viral DNA replication inhibition.

2.1.7 Protein Synthesis Assay

NIH 3T3 cell monolayers were grown to 75% confluency in 12-well plates. The cell sheets were washed with methionine-free medium (Invitrogen), containing 10% NCS, before fresh medium was added to each well. Then 1µCi ^{35}S-methionine was added per 10^5 cells. Monolayers were incubated for 6h at 37°C in a humidified incubator at 5% CO_2.

After this period, the medium was removed and monolayers washed with PBS. The wells were drained and the 12-well plates placed on ice. Ice-cold NP-40 lysis buffer (0.25ml; 150mM NaCl, 1.0% NP-40, 50mM Tris pH8.0) was added to each well and incubated on ice for 30min with gentle agitation to lyse the cells.

Any residual cell layer was then scraped off and the lysate transferred into 1.5ml tubes and vortexed for 1min. The cell debris was removed by centrifugation at 10,000g for 10min and the supernatant transferred into fresh tubes. An aliquot of 5µl for each supernatant was then spotted onto glass fibre filters in at least triplicate for each sample. Filters were then placed into shallow trays that were carefully flooded with ice-cold 5% trichloroacetic acid (TCA). These were incubated on ice for 30min to allow precipitation of protein. The TCA was then drained and replaced before incubation at room temperature for 5min with gentle agitation. This was repeated twice and then the TCA was replaced with 95% ethanol and incubated again at room temperature with gentle agitation for 5min. Filters were then removed and left on foil to dry.

Dry filters were added to 4ml of scintillant (HiSafe, Wallac) in counting vials and incorporation of labelled-methionine was then quantified (Beckman scintillation counter), counting for 1min intervals.

2.2 Virus

2.2.1 Virus Strains

MHV-68 strain g2.4 was kindly provided by Dr Stacy Efstathiou (Division of Virology, Department of Pathology, University of Cambridge). Strain g2.4 was originally isolated as part of the field work on arboviruses in small rodents that was carried out in Slovkia, November 1976, by Blaškovič et al (1980).

2.2.2 Virus Stocks

MHV-68 was propagated in the BHK cell line. Virus stocks were prepared by infecting 95% confluent cell monolayers at moi 0.001. The viral inoculum was prepared in the minimal volume of maintenance medium appropriate for the monolayer, i.e. 20ml for 850cm^2 roller bottles, 5ml for 175cm^2 flasks. Virus was allowed to absorb for 1h at 37°C, with occasional agitation. The inoculum was then discarded, the monolayer washed and then replaced, with fresh maintenance medium (100ml for 850cm^2 roller bottles, 40ml for 175cm^2 flasks). Cells were incubated and harvested when cytopathic effect (cpe) was at least 95% (approximately 5 days).

The infected cells were scraped off and pelleted at 1,600g for 30min at 4°C, in an angled rotor (Beckman JA-21). The supernatent was carefully removed and extracellular virus pelleted at 16,000g for 90min at 4°C, in an angled rotor (Beckman JA-21). The extracellular virus was resuspended in the minimal volume of PBS (25µl PBS per 1 x 10^7 cells), aliquoted and stored at -80°C. If necessary, PBS was added to the viral pellet and left for 16h at 4°C to aid resuspension. Titres of around 10^9 pfu/ml were routinely achieved.

The cellular pellet was resuspended in 200µl of PBS per 1 x 10^7 cells and sonicated at 0°C to free intracellular virus. Cellular debris was removed by centrifugation, at 1,300g for 5min, and the cellular virus-containing supernatent was aliquoted and stored at –80°C. Intracellular virus was used for producing further viral stocks.

2.2.3 Virus Purification

Cell monolayers were infected and virus harvested as previously described (2.2.2) with the exception that the extracellular virus was resuspended in double the volume of PBS (50µl PBS per 1 x 10^7 cells).

The extracellular virus was purified by sucrose density centrifugation. Discontinuous sucrose gradients (14ml of 10-40% sucrose) were prepared in 18ml ultra-clear centrifugation tubes (Beckman). Two ml of each sucrose solution (40-10% in 5% steps) was layered carefully into each tube and then left at 4°C for 16h.

Up to 3ml of extracellular virus was loaded onto each gradient and the weight of each tube adjusted with additional PBS. Tubes were centrifuged at 20,000g for 1h at 4°C under vacuum in an Ultracentrifuge (Beckman; swing-out rotor SW 28.1). Viral bands were visualised by shining a narrow beam of light through the tubes at an angle. Two slightly blue, opaque bands were seen corresponding to the light (upper band) or heavy (lower band) particles. Bands were harvested by inserting a needle through the side of the tube, starting with the upper band. Purified virus was then pelleted at 83,000g for 1h at 4°C before being resuspended in 2.5µl PBS per 1x 10^7 cells, aliquoted and stored at –80°C.

2.2.4 Titration of Virus

Cellular, extracellular and purified virus titres were determined by plaque assay. Ten-fold serial virus dilutions were prepared in maintenance medium and 100µl of the appropriate dilution added to NIH 3T3 cell monolayers, grown in 6-well plates. Virus was adsorbed to the cells at 37°C for 1h with occasional agitation. Monolayers were then rinsed with fresh maintenance medium before overlaying with 2ml

maintenance medium containing 0.8% agarose (Invitrogen). Plates were incubated at 37°C in a humidified 5% CO_2 incubator until plaques were formed (3-5 days). Cells were then fixed with formol saline (4% formaldehyde, 0.85% $NaCl_2$) for 30min and the agarose overlay removed. Cell monolayers were stained with 0.1% crystal violet solution (dissolved in a 1:1 mixture of methanol and ddH_2O) and plaques were counted.

2.2.5 *In Vitro* Infection with Virus

Viral inoculum was prepared by serial dilution of virus stocks. Cell monolayers were grown to confluency and medium removed. Monolayers were then washed with pre-warmed maintenance medium. Monolayers were inoculated with virus in the minimal volume of maintenance medium (e.g. 100µl for 6-well plates), and left to adsorb for 1h at 37°C, in a 5% CO_2 humidified incubator with occasional agitation. Monolayers were washed and overlaid with warm maintenance medium, and then incubated. Mock-infections of monolayers were also set-up, replacing viral inoculum with the equivalent volume of maintenance medium.

2.3 Molecular Biology

2.3.1 Viral DNA Isolation

Virus (1ml of viral stocks) was digested with DNAse I (Roche; 2.3.12.b) and then washed with PBS by 3 rounds of pelleting (3000*g*, 5min) and resuspending. The virus was then pelleted again and resuspended in 1ml proteinase K digestion buffer (0.01M Tris pH7.8, 0.005M EDTA, 0.5% SDS) and proteinase K added to a final concentration of 400µg/ml. The mixture was vortexed and incubated at 50°C for 3h.

An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) saturated with 10mM Tris pH8.0 and 1mM EDTA (Sigma) was added and mixed gently on a rotator for 10min and then a further 20min if an emulsion had not formed. Phases were separated at 3,000g for 10min. The aqueous phase was removed and an equal volume of chloroform was added, and the tube vortexed, before being centrifuged as above. DNA was precipitated from the aqueous phase with 0.1 volumes 3M sodium acetate (NaOAc) pH5.2 and 2 volumes of ice-cold ethanol, mixed gently and pelleted at 5,000g for 5min. The pellet was washed with 1ml 70% ethanol and then the ethanol was removed. Residual ethanol was left to evaporate, before the DNA was resuspended in Tris-EDTA buffer (TE buffer). DNA was redissolved on ice for 15min with occasional gentle mixing. Integrity of the DNA was

checked by 0.8% agarose gel electrophoresis (2.3.2), before aliquoting and storage at -20°C.

2.3.2 Agarose Gel Analysis of DNA

Agarose gels were prepared at a variety of concentrations ranging from 0.8% to 2%, depending on the size of the nucleic acid being analysed. Agarose (Ultrapure electrophoresis grade, Invitrogen) was melted in 1x Tris-acetate-EDTA buffer (TAE buffer; 0.04mM Tris, 0.001mM EDTA, 0.02M acetic acid) and 2µl/100ml ethidium bromide (Sigma) added. The gel was set in a gel rig with the appropriate comb and then they were submerged in 1x TAE buffer. Samples (usually 5µl) were mixed with 6x loading buffer (Sigma) prior to loading. DNA ladders (5µl of 1kb+ DNA marker, Invitrogen) were used as reference markers for the electrophoresis, unless otherwise indicated. DNA was electrophoresed and then visualised under UV and photographed using Gelworks software and camera (UVI).

2.3.3 Isolation of DNA from Agarose Gels

DNA bands were excised from ethidium bromide-stained agarose gels using sterile scalpels under UV. DNA was purified from the bands using the QIAquick gel purification kit (Qiagen) following the manufacturer's instructions. This kit uses a silica gel membrane to capture the DNA.

2.3.4 Spectrophotometric Analysis of Nucleic Acids

Samples for spectrophotometric analysis were diluted in TE buffer to an appropriate volume and compared to a blank sample consisting of TE buffer alone. The concentration of nucleic acid was estimated as 1 absorbance unit at 260nm (A_{260nm}) being equivalent to 50µg/ml dsDNA, 37µg/ml ssDNA and 40µg/ml RNA. The ratio of absorbance at 260nm and 280nm (A_{260nm}:A_{280nm}) was also calculated, as ratios of 1.8-2.2 indicate pure DNA/RNA with minimal protein contamination.

2.3.5 Restriction Endonuclease Digestion of DNA

Buffer systems and restriction endonucleases (Roche) were used according to the manufacturer's recommendations. Generally, reactions were performed in 10µl volumes at 37°C for 1h, with 10U of enzyme and up to 10µg of DNA.

2.3.6 Polymerase Chain Reaction

Various PCR protocols were used depending on the application. Generally PCR was performed using 40U/ml Taq polymerase (Invitrogen), 1x PCR buffer

(Invitrogen), 1.5mM magnesium chloride (Invitrogen), 0.2mM deoxynucleoside triphosphates (dNTPs; Promega), 1μM each primer pair, and template in a final volume of 25μl, using master mixes whenever possible. (Using master mixes involved preparing the PCR reagents for all reactions together before aliquoting out individual reactions and adding template.) The PCR cycling parameters were 5min at 95°C, followed by 30 cycles of 95°C (30s), 58°C (30s) and 72°C (30s), before a final elongation of 5min at 72°C and holding at 4°C. The products of the reaction were analysed by agarose gel electrophoresis (2.3.2).

2.3.7 Isolation of DNA from PCR Reactions

DNA was isolated from PCR reaction mixes using the QIAquick PCR purification kit (Qiagen) following the manufacturer's instructions. DNA was captured from the reaction mixture by a silica gel membrane. Affinity of the DNA for the membrane was controlled by changing the ionic interactions between the membrane and the DNA.

2.3.8 Cloning of DNA Fragments

The pGEM-T Easy vector system (Promega) was used for all cloning procedures following the manufacturer's protocol. This system employs a plasmid vector with a multiple cloning site within a beta galactosidase (lacZ) cassette, resulting in blue and white colonies, white indicating an insert. DNA fragments to be cloned were gel-purified (2.3.3) and then ligated at room temperature for 1h.

The transformed cultures were plated onto duplicate L-agar plates (Imperial) containing 100μg/ml ampicillin, and which had been spread with 200μl of 100mM isopropyl β-D-thiogalactopyranoside (IPTG) and 20μl of 50mg/ml 5-bromo-4-chloro-3-indolyl-β-D- galactopyranoside (X-gal) and left to absorb for 30min at 37°C prior to use. White bacterial colonies contained plasmids with successfully ligated inserts and were therefore selected over blue colonies.

Following isolation of plasmid DNA (2.3.9), the DNA was digested with the appropriate endonucleases (2.3.5) to check for the presence and orientation of any insert. One of the chosen endonucleases had a restriction site complementary to the plasmid, and the other had one matching the insert sequence. The size of the digested fragment would then indicate the orientation of the insert. A second digest was performed with a restriction enzyme cutting either side of the insert to show its size.

2.3.9 Small Scale Preparation of Plasmid DNA

Single white colonies were picked from plates and used to inoculate 5ml L-broth (Imperial) containing 100μg/ml ampicillin. Cultures were incubated for 16h at 37°C with continuous vigorous shaking (180rpm). Each single colony was also dipped into a PCR reaction mix to be used as template for a colony PCR (2.3.6), which was used to confirm the identity of the insert. Briefly, the primers used in the colony PCR reaction were specific for the expected insert and thus successful PCR indicated that the correct insert was present. The overnight cultures were collected the next day by centrifugation and the QIAprep kit (Qiagen) used for plasmid isolation.

2.3.10 Bacterial Stocks

Bacterial stocks were prepared by freezing cultures in 25% glycerol, 75% L-broth (Imperial) at -80°C. To reculture the bacterial stock, a sterile pipette tip was stabbed into the frozen stock and used to inoculate L-broth containing ampicillin.

2.3.11 Sequencing

2.3.11.a Manual Sequencing

Plasmid DNA was isolated from 5ml bacterial cultures using the QIAprep kit (Qiagen; 2.3.9) and resuspended in 32μl H_2O. The plasmid DNA was then prepared according to the manufacturer's protocol (T7 sequenase v2.0 7-deaza-dGTP sequencing kit; Amersham). Briefly, plasmid DNA was denatured by adding NaOH to 200mM and EDTA to 20mM. This was mixed and incubated at room temperature for 5min. The DNA was then precipitated by adding 0.1 volumes 3M sodium acetate pH5.2 and 2.5 volumes ice-cold ethanol, and then incubating at -70°C, for at least 30min. DNA was pelleted for 25min at 13,000*g* at 4°C, and then washed with 70% ethanol before the pellet was dried in a vacuum concentrator.

The manufacturer's protocol was also followed to perform the actual dideoxy sequencing reactions.

Sequencing gel plates were thoroughly cleaned with ethanol and the smaller plate siliconised with Sigmacote (Sigma). Following assembly of the plates, 6x Gel-Mix (Invitrogen) was used to produce the polyacrylamide sequencing gel following the manufacturer's instructions. Briefly, gel polymerisation was initiated with the addition of ammonium persulphate and a 0.4mm gel poured between the sequencing gel plates. Following 40min at room temperature to allow the gel to set, the glass plate spacers were removed and the gel prerun in a vertical sequencing

gel tank (Invitrogen) for 1h at 50-70W to equilibrate the gel with TBE buffer (0.089M Tris, 0.089M Boric acid, 0.002M EDTA).

Once the samples had been prepared as per the manufacturers instructions, they were denatured at 80°C for 2min and half of each sample was loaded into wells to be electrophoresed at 50-70W for 2h or until the loading dye's frontline had run out completely. The remaining half of each sample was then denatured and loaded into adjacent lanes on the gel, for the second run. Therefore, each sample was electrophoresed twice, a long 4h run and a short 2h run.

Once finished the glass plates were carefully prised apart and 3MM paper (Whatman) used to remove the gel. The exposed side was covered in cling film and the gel dried under vacuum at 80°C for 1h before being exposed to film.

2.3.11.b Automated Sequencing

Automated sequencing was performed using the Beckman Coulter CEQ2000XL Sequencer and CEQ Dye Terminator Cycle Sequencing kit (Beckman) following the manufacturers' recommendations. Analysis of the resulting output was performed using ABI Prism software.

2.3.12 Total RNA Isolation

Special care to avoid RNAse contamination was taken for all procedures involving RNA. Separate chemical solutions for RNA work were prepared to minimise RNAse contamination. These were prepared from separate chemical stocks and diethyl pyrocarbonate-treated water (DEPC-H_2O). All apparatus used for RNA work was kept separate where possible, and cleaned regularly with DEPC-H_2O. RNAse-free DEPC-H_2O was prepared by adding 1ml DEPC to 1l H_2O. This mixture was shaken vigorously to dissolve the DEPC and then left to stand for 16h. Finally the DEPC-H_2O was autoclaved before use.

This RNA isolation protocol consisted of several steps: initial RNA isolation, DNAse treatment and re-isolation of RNA.

2.3.12.a RNA Isolation

Tissue culture samples were prepared by removing media and adding 1 ml Trizol per 10^6 cells directly. Samples were then vortexed vigorously for 1min before being left to stand on ice for 5min. Chloroform (0.2ml) was added and vortexed for 1min, before being placed on ice for 2min. Phases were separated by centrifugation at 15,000g for 10min at 4°C. The upper aqueous phase was transferred to a fresh

tube, being sure to avoid the interphase and RNA precipitated with an equal volume of isopropanol for 10min on ice and then pelleted at 15,000g for 30min at 4°C. RNA was washed with 1ml 80% ethanol and then repelleted at 15,000g for 5min at 4°C, before resuspending the RNA in 88µl DEPC-H_2O.

2.3.12.b DNAse I Treatment

DNAse buffer (10µl of 10x stocks: 400mM Tris-HCl pH7.5, 100mM $NaCl_2$, 60mM $MgCl_2$) and 2µl DNAse I (10U/µl, Roche) were added to the sample and incubated at 37°C for 30min, at which time 10µl 0.1M EDTA pH8.0 was added to terminate the reaction.

2.3.12.c Phenol Chloroform Extraction

The RNA solution was made up to 500µl with DEPC-H_2O and phenol:chloroform extracted: An equal volume of phenol:chloroform 5:1 pH4.7 (Sigma) was added and then vortexed for 1min, before incubation on ice for 5min. A further 0.2ml chloroform was then added, vortexed for 1min, and placed on ice for 2min. Phases were then separated by centrifugation at 15,000*g* for 10min at 4°C.

The upper aqueous phase was transferred to a fresh tube, being sure to avoid the interphase and RNA precipitated with 0.1 volumes 3M NaOAc pH5.2 and 2.5 volumes ice-cold ethanol at –80°C for at least 2h. RNA was pelleted at 15,000*g* for 30min at 4°C, washed with 1ml 80% ethanol and then repelleted at 15,000*g* for 5min at 4°C, before resuspending the RNA in an appropriate volume of DEPC-H_2O (usually 10-20µl).

2.3.13 Denaturing Gel Analysis of RNA

All RNA was tested for quality by electrophoresis on a denaturing gel and by spectrophotometry (2.3.4).

Denaturing agarose gels were prepared by melting 1g agarose in 74ml H_2O and 10ml 10x MOPS buffer (0.2M MOPS, 0.05M, 0.001M EDTA pH7.0). This was allowed to cool to room temperature before addition of 16ml 37% formaldehyde. The gel was then poured into a gel rig and left to set, at which stage 1x MOPS buffer was added as running buffer.

Two volumes of 1.5x RNA loading buffer (1.5x loading dye (0.06% w/v bromophenol blue, 0.06% w/v xylene cyanol), 1.5x MOPS buffer, 9% formaldehyde, 60% deionised formamine, 5% ethidium bromide) was added to 5ml of each RNA sample, and samples denatured at 65°C for 15min before chilling on ice.

The gel was prerun for 5min at 20V before samples were loaded, and then run until the dye had reached approximately two-thirds down the gel. RNA was visualised on a UV lightbox.

2.3.14 Analysis of RNA Using RNA LabChips

The manufacturer's protocol was followed to setup and perform RNA analysis on RNA LabChips using a Bioanalyser 2001 (Agilent). Briefly, the fluorescent nucleotide dye was mixed with the gel matrix, before being injected into the micro capillaries of the RNA LabChip under pressure. The RNA samples were then loaded onto the LabChip, along with the supplied RNA marker. The RNA was electrophoresed through the microcapillary channels and fluorescence at a set point, measured as a function of time. Electropherograms were generated for each sample and analysed for signs of RNA degradation.

2.3.15 Northern Blot Analysis

2.3.15.a Northern Blotting

RNA was electrophoresed on a denaturing gel (2.3.13), ensuring that the same amount of RNA was loaded in each lane. RNA was visualised briefly on a UV lightbox and a picture taken of the gel with a ruler lined up for future reference.

The gel was placed on 3MM paper (Whatman) as shown in Fig. 2.2 and nylon membrane (Hybond N+, Amersham) positioned to line up with the wells of the gel. The top left corner was cut to aid orientation. Cling film was used on the edges of the membrane to prevent shorting of the circuit (i.e. to ensure that no liquid bypassed the nylon membrane). The rest of the apparatus was assembled as shown and left for 16h.

The membrane was then lifted off avoiding contact with possible sources of nucleic acids and the RNA covalently bonded to the nylon by UV irradiation (Strategene).

2.3.15.b Synthesis of Probe for Northern Blot Analysis

An appropriate sequence was selected to be the probe sequence, usually 100-500bp in size, and the appropriate primers synthesised commercially (MWG Biotech). These primers were used to amplify the probe sequence by PCR from a sequence-verified plasmid containing the appropriate DNA insert, and then gel purified (2.3.3). The DNA was quantified by electrophoresis on an agarose gel along with a quantitative ladder (Bioline).

The High Prime DNA labelling kit (Roche) was used to produce a ^{32}P-labelled probe, following the manufacturer's instructions. The labelled probe was purified using the Nucleospin kit (a silica gel membrane based capture column; Clontech).

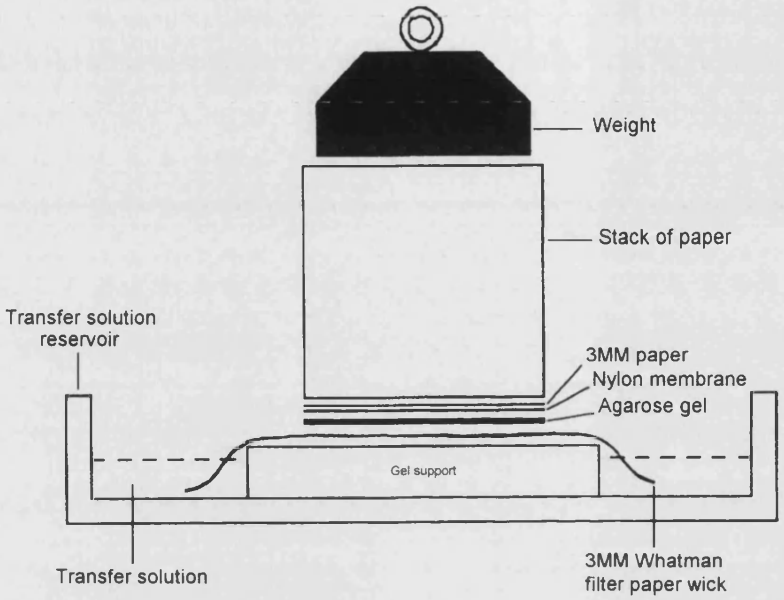


Figure 2.1. Diagrammatic representation of apparatus involved in setting up a northern blot.

2.3.15.c Hybridisation and Washing

Membranes were prehybridised in 10ml hybridisation solution (Expresshyb, Clontech) containing 1mg/ml denatured salmon sperm DNA (Sigma) at 65°C for 1h in a rotating incubator (Hybaid). The hybridisation solution was then discarded and replaced, and 1mg/ml denatured salmon sperm DNA and probe added for hybridisation at 65°C for 1h. The membrane was then washed 3 times in 10ml 2x saline-sodium citrate (SSC) buffer (0.03M sodium citrate pH7.0, 0.3M sodium chloride), 1% sodium dodecyl sulphate (SDS) for 15min each at room temperature. The membrane was then washed again in 10ml 0.1X SSC buffer, 0.5% SDS, 3 times for 15min each at 58°C. Finally the membrane was wrapped in cling film (being careful to avoid bubbles) and visualised by autoradiography or using a phosphor screen and imager.

2.3.16 cDNA Synthesis

Typically for reverse transcription (RT), 2-5μg total RNA, 1x first strand buffer (50mM Tris-HCl pH8.3, 75mM potassium chloride, 3mM magnesium chloride; Invitrogen), 1mM dithiothreitol (DTT; Invitrogen), 1mM dNTPs, 0.25μM primers and 2.5U/μl Mouse Moloney murine leukemia virus (MMLV) reverse transcriptase (RT; Invitrogen) were combined in a total volume of 20μl of DEPC-H_2O, and incubated at 42°C for 15min. The reaction was then terminated by 5min incubation at 99°C and then cooled to 4°C.

If specific primers were to be used for reverse transcription, then the RNA and primers were denatured and annealed by heating to 70°C for 5min and then placing on ice. If anchored oligo-dT primer was used then the complete reaction mix was incubated at room temperature for 10min prior to the incubation step.

2.3.17 Stripping Membranes for Re-hybridisation

If blots were to be used again, they were stripped: 500 ml 0.5% SDS was boiled and then poured into a plastic container with lid. The membranes were immersed in the SDS solution and agitated for 1h. Each stripped membrane was then checked for any remaining hybridised probe, using a phosphor screen and imager. If residual signal was detected, the stripping process was repeated. Care was taken to prevent the blots from drying out.

2.3.18 Differential Display

2.3.18.a cDNA Synthesis

Two μg of RNA was mixed with anchored oligo-dT at 0.2μM in a volume of 5μl DEPC-H_2O and spun briefly. The mixture was heated to 70°C for 3min and then placed on ice for at least 2min before a short centrifugation.

1x first strand buffer (Invitrogen), 1mM dNTPs and 5U/μl MMLV RT (Invitrogen) were added to the RNA and primer mixture to make a total volume of 10μl, using master mixes when possible. This was mixed and centrifuged briefly before incubating at 42°C for 1h, 75°C for 10min and then placed on ice.

Each reaction was split into two and diluted: 8μl into 80μl water (cDNA A) and 2μl into 80μl water (cDNA B). Samples could be stored at –20°C at this stage.

2.3.18.b Differential Display PCR

The differential display PCR (ddPCR) protocol used a number of degenerate primer pairs to generate DNA 'fingerprints'. Combinations of 5 primers were used initially to test the system.

Both dilutions of cDNA were used with each differential display primer combination, as was a negative H_2O control (H_2O instead of cDNA). Furthermore, a number of negative RNA controls (RNA instead of cDNA) were also run. This resulted in a large number of reactions and therefore tables such as Table 2.2 were employed to avoid confusion.

One µl of the cDNA sample was mixed with 1µl each of the two primers to be used. The cDNA was substituted with H_2O and RNA as appropriate for the controls.

A master mix for the reaction was prepared by combining KlenTaq PCR reaction buffer (40mM Tricine-KOH pH9.2 at 25°C, 15mM KOAc, 3.5mM MgOAc, 3.75µg/ml bovine serum albumin; Clontech), 50µM dNTPs, Advantage Klentaq polymerase and hotstart antibody mix (1% glycerol, 0.8mM Tris-HCl pH7.5, 1.0mM KCl, 0.5mM $(NH_4)_2SO_4$, 2.0mM EDTA, 0.1mM β-mercaptoethanol, 0.005% Thesit; Clontech), 50nM [α-^{33}P]dATP (1000-3000 Ci/mmol; Amersham) and ddH_2O in a final volume of 17µl per reaction. The master mix was vortexed and spun briefly before adding to the cDNA and primers.

The reaction mixes were cycled as follows:

95°C	5min		
40°C	5min		
68°C	5min		
94°C	30s	}	
40°C	30s	}	2 cycles
68°C	5min	}	
94°C	20s	}	
60°C	30s	}	23 cycles
68°C	2min	}	
68°C	7min		
4°C	hold		

Following the ddPCR, reaction mixes could be stored at –20°C.

Table 2.1 Sequence of primers used for ddPCR

Name	Sequence (5'-3')
P1	ATTAACCCTCACTAAATGCTGGGGA
P2	ATTAACCCTCACTAAATCGGTCATAG
P3	ATTAACCCTCACTAAATGCTGGTGG
T8	CATTATGCTGAGTGATATCTTTTTTTTGC
T9	CATTATGCTGAGTGATATCTTTTTTTTGG

Table 2.2 Sample organisational and labelling chart for setting up ddPCR

Tube Label	cDNA Sample	Primers	Tube Label	cDNA Sample	Primers
Experimental displays			Water controls for each primer pair		
1	1A	P1 & T9	H1	Water	P1 & T9
2	1B	P1 & T9	H2	Water	P1 & T8
3	2A	P1 & T9	H3	Water	P2 & T9
4	2B	P1 & T9	Etc.		
5	1A	P1 &T8	Total RNA controls for each RNA		
6	1B	P1 &T8	(Use any of above primer pairs)		
7	2A	P1 &T8	R1	RNA 1	P1 & T9
8	2B	P1 &T8	R2	RNA 2	P1 & T9
			Etc.		
9	1A	P2 & T9			
10	1B	P2 & T9			
11	2A	P2 & T9			
12	2B	P2 & T9			
Etc.					

2.3.18.c Electrophoresis and Autoradiography

A denaturing polyacrylamide urea gel was prepared and prerun (2.3.11.a). Five µl of each PCR reaction was mixed with 5µl of ddPCR loading buffer (95% formamide, 0.2% bromophenol blue, 0.2% xylene cyanol), denatured at 95°C, 2min and then placed on ice. Two µl of each reaction was loaded and the gel run at 70W for 2h or until the xylene had run out of the gel. Each reaction was then loaded again and the gel run for another 2h resulting in a long and short run. The gel was dried (2.3.11.a) and exposed to a phosphor screen (Fuji BAS Image Analysis System). After reading on the phosphoimager, the gel was carefully aligned and used to expose a

film. Three corners of the gel and film were cut to aid location of bands of interest on the gel.

2.3.18.d Purification of DNA Fragments from Dried Acrylamide Gels

Following analysis of the differential display, bands of interest were identified and located on the autoradiography film. The film was then lined up with the dried gel and the two taped together. The positions of bands identified on the film were marked on the gel by pushing a sharp pin through both surfaces. The film was then removed and bands cut out of the gel using a fresh scalpel for each one. Care was taken to cut out gel fragments without taking any paper. Gel fragments were transferred to 0.5ml tubes containing 40µl of TE buffer. These were heated at 100°C for 5min to melt the agarose and release the DNA into the TE buffer. At this stage samples could be stored at –20°C.

Gels were re-exposed and visualised to confirm that the correct bands had been excised.

2.3.18.e Reamplification of Bands of Interest

Seven µl of the DNA solution (2.3.18.d) was used as a template for PCR to amplify the band of interest. The DNA solution was mixed with KlenTaq PCR reaction buffer (2.3.18.b), 50µM dNTPs, Advantage KlenTaq polymerase and hotstart antibody mix (2.3.18.b), 1µM of each primer that was used originally to produce the band and sterile water in a final volume of 50µl.

This reaction mix was cycled at 94°C for 1min, 60°C for 1min and 68°C for 2min for a total of 20 cycles, before 10µl was analysed by electrophoresis through a 2% agarose gel.

2.3.18.f Northern Blot Confirmation of ddPCR

The reamplified band of interest was excised and purified from the gel (2.3.3) and then cloned into a plasmid (2.3.8). This plasmid was then purified (2.3.9) and sequenced (2.3.11) to identify the band of interest. Finally, the plasmid was used to prepare probe for a northern blot analysis (2.3.15) to confirm the differential expression seen with the ddPCR.

2.3.19 Preparation of Probes for DNA Arrays

Primers were designed to PCR amplify approximately 300bp fragments of DNA from the 5' end of each predicted MHV-68 ORF. In addition to the predicted ORFs,

inter-gene regions of MHV-68's genome of greater than 100bp were also included. If the inter-genic regions exceeded 300bp (e.g. approximately 840bp between ORF K3 and M5) then two cDNA were amplified to correspond to the 5' end of a potential transcript that could be transcribed in either direction. Appendix I list the primers used for generation of probe cDNA sequences.

The following genes were chosen as positive and negative hybridisation controls: Glyceraldehyde 3-phosphate dehydrogenase (GAPDH; accession no. NM008084), myosin 1 (L00923), murine ornithine decarboxylase (MOD; M10624), β-actin (X03672), calcium binding protein (Cab) 45 (U45977), ribosomal protein S29 (NM009093), ubiquitin (AF285162), phospholipase A2 (D78647), hypoxanthinie phosphoribosyl transferase (HPRT; J00423), pBluescript II (SK+) plasmid (Stratagene), tobacco mosaic virus (TMV) 180kDa protein (D78608) and water. Luciferase (E15166) was included as an internal control for normalisation of arrays.

All PCR products were cloned into pGEM-T Easy vectors (2.3.8) and the presence of the correct insert verified by PCR (2.3.6) with insert-specific primers, and the presence of a single insert verified by PCR with pGEM-T Easy vector-specific primers (5'-CCA TGG CGG CCG CGG GAA TT-3', 5'-GGC GGC CGC GAA TTC ACT AG-3'). Finally, the sequence of the DNA insert was verified (2.3.11.b) to ensure that there had been no misincorporation of bases during the PCR amplification.

DNA species were then reamplified from the relevant plasmids using a high yield PCR protocol: 1x PCR buffer (Invitrogen), 1.5mM $MgCl_2$, 2mM dNTPs (Promega), 4μM pGEM-T Easy vector-specific primers, 50U/ml Taq polymerase (Invitrogen) and 30pg/μl template in a final volume of 100μl, using master mixes whenever possible. This reaction was incubated for 5min at 95°C, followed by 30 cycles of 95°C (30s), 55°C (30s) and 72°C (30s), before a final elongation of 5min at 72°C, and then placed on ice. The amplified probes were purified (2.3.3) and then analysed by spectrophotometry (2.3.4). This protocol returned yields of approximately 5μg per reaction.

2.3.20 Vacuum-Spotting DNA Arrays

DNA (800ng) was made up to 100μl with Tris buffer (10mM Tris-HCl pH8.5) and then denatured at 95°C for 5min, before placing on ice. Each sample was denatured with 1vol 20x SSC.

The Bio-Dot microfiltration apparatus (Bio-Rad) was assembled according to the manufacturer's instructions using Hybond N+ nylon membrane (Amersham) prewetted in 10x SSC, and 2 sheets of Bio-Dot filter paper (Bio-Rad), also prewetted. A corner of the membrane was cut to aid orientation. Briefly, 200µl of 10x SSC was drawn through the wells under a gentle vacuum twice, and then followed by the actual samples. A further 200µl of 10x SCC was drawn through and then the apparatus was disassembled and the membrane placed on several sheets of filter paper soaked in denaturing solution (1.5M NaCl, 0.5M NaOH) for 5min. The membrane was then transferred to several sheets of filter paper soaked in neutralising solution (1.5M NaCl, 0.5M Tris-HCl pH7.2, 1mM EDTA) for 5min, before drying and cross-linking (Stratagene). Arrays were stored in cling-film at 4°C. The layout of these vacuum-spotted arrays is shown in Appendix II.

2.3.21 Automated Manufacture of DNA Arrays

The Beckman Biomek 2000 robot was programmed to print arrays using the integrated tools. Each probe was spotted in triplicate at 4 concentrations: 100ng/µl, 50ng/µl, 25ng/µl, 12.5ng/µl, onto Hybond N+ nylon membranes (Amersham). The layout of these Biomek arrays is shown in Appendix II.

2.3.22 Manually Spotted DNA Arrays Using the Multiblotter Pin-Tool

Arrays were spotted with 50ng of probe DNA in duplicate onto Hybond N+ nylon membranes (Amersham) using a 384-pin multiblotter (V&P Scientific) and the membranes denatured prior to use (2.3.24). Arrays were stored at room temperature in sealed bags. The layout of these pin-tool arrays is shown in Appendix II.

2.3.23 Automated Manufacture of Oligo Arrays

Oligonucleotide-based arrays (oligo arrays) were manufactured using a robot system (Rosetta and Agilent). The arrays were manufactured by the array group at Arrow Therapeutics, following standardised protocols. Briefly, probes were synthesised *in situ* on glass slides using standard phosphoramidite chemistry. Each feature on the array was individually spatially addressed by 2 ink-jet printer nozzles, one providing a tetrazole activator (necessary for the coupling reaction), and the second providing the appropriate base monomer for that layer's addition. The ancillary reagents for DNA synthesis (oxidiser, deblocking reagent and acetonitrile washes) were delivered by flooding across the array. The layout of these oligo arrays is shown in Appendix II.

2.3.24 Denaturing Probes Spotted on DNA Array Membranes

Arrays with probes that had not been denatured prior to spotting were placed on filter paper soaked with denaturation buffer (0.66M NaCl, 0.5M NaOH) for 10min and then washed in ddH_2O for 10min and finally in 40mM phosphate buffer pH7.3 for 10min. Membranes were then dried and stored in plastic at 4°C.

2.3.25 Preparation of Radiolabelled cDNA Target for DNA Array Analysis

Labelled cDNA target for hybridisation to arrays was prepared by reverse transcription of RNA, whilst incorporating radiolabelled nucleotides. Generally, 10μg RNA (2.3.12; 2.3.13) was spiked with 10ng luciferase RNA (2.3.29), and primed with a mixture of primers complementary to the 3' sequence of the DNA probes (0.2mM; array primer mix). The RNA and primers were annealed by incubating at 70°C for 5min and then placed on ice.

The reverse transcription was carried out in 1x first strand buffer, 1mM DTT, 1mM dT/G/CTP, 2.5μCi/μl [α-^{33}P] dATP (2500 Ci/mmol, Amersham) and 5U/μl Superscript II reverse transcriptase (Invitrogen), combined in a total volume of 20μl. The reaction mixture was incubated at 50°C for 90min, before terminating with 1μl 100mM EDTA and placing on ice.

The labelled target was purified using the Nucleospin kit (Clontech) and denatured at 95°C for 5min prior to hybridisation.

2.3.26 Preparation of Fluorescently-labelled cDNA Target for Oligo Array Analysis

Labelled cDNA target for hybridisation to arrays was prepared by reverse transcription of RNA, whilst incorporating fluorescently labelled nucleotides. Generally, 10μg RNA (2.3.12; 2.3.13) spiked with 10ng luciferase RNA (2.3.29), was primed with random hexamers (0.2mM). RNA and primers were annealed by incubating at 70°C for 5min and then placing on ice.

The reverse transcription was carried out in 1x first strand buffer, 1mM DTT, 1mM dT/G/ATP, 0.05mM Cy5-dATP (NEN) and 5U/μl Superscript II reverse transcriptase, combined in a total volume of 20μl, and incubated at 37°C for 16h. The reaction was then terminated with 1.5 μl 20mM EDTA. The RNA was hydrolysed by adding 15μl freshly prepared 0.1M NaOH and incubating at 70°C for 10min. Finally 15μl of 0.1M HCl was added to neutralise the alkali.

The labelled target was purified using the QIAquick PCR purification kit (Qiagen), following the manufacturer's protocol except for using ddH_2O as the final elutant. The labelled target was denatured at 95°C for 5min prior to hybridisation.

2.3.27 DNA Array Hybridisation

Membranes were prehybridised in the hybridisation solution (Expresshyb, Clontech) containing 1mg/ml denatured salmon sperm DNA and 0.1mg/ml murine C_0T-1 DNA (Invitrogen) at 60°C for at least 30min in a rotating incubator. The hybridisation solution was then discarded and replaced along with fresh denatured salmon sperm DNA and C_0T-1 DNA. Finally the denatured target was added and hybridised to the array at 60°C for 16h. The membrane was then washed 3 times in 2x SSC buffer, 1% SDS for 15min each at room temperature. The membrane was then washed again in 0.1X SSC buffer, 0.5% SDS, 3 times for 15min each at 60°C. Finally the membrane was wrapped in cling film and placed on a phosphor screen.

2.3.28 Oligo Array Hybridisation

The glass arrays were placed into hybridisation chambers and hybridisation solution added (1M NaCl, 1 x MES buffer pH6.4, 20mM EDTA, 20% formamide, 1% TritonX 100). Denatured target was then injected into the chamber. The hybridisation was performed for 16h at 75°C in a rotating incubator. The arrays were then washed in Rosetta buffer 1 (6 x SSPE, 0.005% sarcosine) for 5min and then Rosetta buffer 2 (0.06 x SSPE, 0.18% PEG 200) for 5min. Finally, the arrays were dipped in ether and air-dried with the aid of a compressed air gun. Dry arrays were stored away from light until ready to be scanned.

2.3.29 SP6 RNA Polymerase Amplification of Luciferase RNA

The SP6 Ribomax RNA amplification kit (Promega) was used to produce luciferase RNA, following the manufacturer's protocol. Briefly, linear dsDNA encoding a SP6 RNA polymerase promotor and the luciferase gene (Promega) was used as template. The template was added to the reaction buffer, ribonucleotides and T7 RNA polymerase, and incubated at 37°C for 4h. The DNA template was then removed by digestion with DNAse I. The RNA was purified by a standard phenol chloroform extraction, and then precipitated with NaOAc and ethanol (2.3.12.c). The precipitation step was performed 3 times to remove unincorporated nucleotides.

The resulting luciferase was electrophoresed through a denaturing RNA gel (2.3.13) to confirm that a single sized RNA species was present and then quantified by

spectrophotometry (2.3.4). The luciferase RNA was stored in 1mg/ml aliquots at -80°C.

2.4 Bioinformatics and Data Analysis

2.4.1 Statistical Tests and Measures

Statistical analysis was routinely employed on all datasets that were of sufficient size. Commonly applied tests and measures included calculating the correlation coefficient, standard deviation and coefficient of variation.

2.4.2 Basic Local Alignment Search Tool

Sequence similarity searches were performed using programs employing the BLAST (basic local alignment search tool) algorithm (Altschul, 1990), if query sequences were more than 100bp in length.

2.4.3 FastA

The FastA program, version 34t10d3 (Pearson, 1988), was used to compare sequences to a database of reference sequences. Alignments were forced to be based on the Smith-Waterman algorithm (Smith, 1981), and default parameters used for all other options.

2.4.3.a Cluster

Gene expression profiles were transformed onto a relative scale so that peak transcript abundance = 1, and then onto a logarithmic scale by taking the square root of each data point. This transformed data was analysed by pairwise comparisons using the Cluster algorithm (Eisen, 1998). The program was configured to perform hierarchical clustering using the average-linkage algorithm and an uncentred correlation matrix. The results of these analyses were visualised with Treeview software (Eisen, 1998).

2.4.4 Design of Oligonucleotide Primers for PCR

Primers were designed using the Primer3 software developed by the Whitehead Institute for Biomedical Research (Rozen, 2000). Primers for PCR were designed using the following parameters:

- Primer size – minimum 18bp, optimum 20bp, maximum 22bp.
- Primer melting temperature – minimum 58°C, optimum 60°C, maximum 62°C.
- Primer GC content - minimum 40%, optimum 50%, maximum 60%.

- 3' GC clamp – 2bp
- All other variables were left at default values.

2.4.5 Design of the Oligo Array Probeset

The MHV-68 genome was split into 100bp sections and one 60-mer oligonucleotide probe was designed for each 100bp section using Hotspots software (Arrow Therapeutics, London). Where possible, the oligonucleotide probes were designed to match up with the DNA probes used for the DNA array so that comparisons between the 2 platforms could be made at a later stage. The Hotspots software was used to analyse each 100bp section and identify regions of low cross-hybridisation potential. The software then calculated the thermodynamic characteristics of every probe that could potentially represent these regions. Probes were then selected by the software based on thermodynamic profile best matches to the hybridisation conditions and to the other probes on the array. Probes exhibiting self-complementarity and defined 'tip' sequences that produce aberrant hybridisation results were excluded by the software . This part of designing the probeset was performed by the array group at Arrow Therapeutics.

To check the validity of probes designed by the custom software, all probes were aligned against the whole genome sequence for MHV-68 using the FastA algorithm to ensure that each probe matched MHV-68's genomic sequence (2.4.3). The probes were also aligned against MHV-68 ORF sequences and the DNA array probe sequences.

Any probes that did not match the genomic sequence uniquely and perfectly were removed from the probeset. Any probes found to extend beyond the limits of an ORF were also discarded. Finally, any probes that showed complementarity to more than one ORF, were discarded as well.

All probes that matched with the DNA array probe sequences were highlighted for later comparison to DNA array data. Probes that were incomplete matches to the DNA array data were also noted.

Each probe was replicated 5 times in the probeset, and then randomly allocated a position on the array. The overall design was inputted to the Rosetta/Agilent robot for manufacture of the arrays (2.3.23). The oligo array probeset is listed in Appendix III.

2.4.6 Quantification of DNA Array Data

Hybridised arrays were used to expose SR phosphor screens (Packard). Screens were scanned with the Cyclone Storage Phosphor System and analysed with Optiquant software (Packard). Quantification of signals on arrays was standardised: scanned images were overlaid with a fixed template, which defined signal regions corresponding to probes on the array. Quantified signal values were imported into spreadsheet software for further analysis. Replicate signals on each array were averaged (mean) and background subtracted. Background was defined as the hybridisation signal value of negative control probes on the array (pBluescript II plasmid and TMV 180kDa protein). All signal values were then normalised to the internal control luciferase to give a signal ratio that was relative to 10ng luciferase mRNA.

2.4.7 Quantification of Oligo Array Data

Each Agilent array was quantified using a DNA microarray scanner with 5μm resolution and 2 lasers, with confocal autofocus (Agilent). The image analysis software on the scanner produced a TIFF file and identified probes on the array image, based on the oligo array design template used to synthesise the probes. The software then removed background effects and identified features that were outliers and could therefore possibly distort the data. Finally, the software normalised the data for different dye incorporation rates and then calculated a differential gene expression value based on Cy5:Cy3 intensities if double-labelling was used, or returned absolute values if single colour labelling was used. The numerical values for each probe along with supporting statistical data were exported into spreadsheet software for further analysis.

3 Differential Display

3.1 The Differential Display System

The differential display system is a powerful technique first developed by Liang and Pardee, that allows the comparison of two RNA populations, by highlighting any differences between them (Liang, 1992). In the instance of a viral infection, a differential display of the RNA population can highlight viral RNAs, as well as any host genes that are regulated by the virus. This provides a complete transcriptional representation of the effects of a viral infection in a host cell.

The aim of this study was to design and optimise a differential display system to examine MHV-68 infection *in vitro*. The changes in host and virus gene expression that occur during infection could then be analysed using the differential display system.

The original protocol (Liang, 1992) and various commercially available protocols (Clontech, GenHunter, Beckman Coulter) were examined. Of these, one protocol (Delta Differential Display System, Clontech) was chosen as the template protocol for this study, as it incorporates improvements made by various groups (Ralph, 1993; Bauer, 1993; Wong, 1994). An overview of the protocol is shown in Figure 3.1.

3.2 Preparation of RNA Template

High quality RNA template is essential for the differential display system as all subsequent steps are dependent on it. Two methods for isolation of RNA were assessed: A phenol:chloroform isolation (TRIzol, Invitrogen), and a silica gel isolation (RNAeasy, Qiagen). Both methods use a guanidium isothiocyanate denaturation step to release the RNA from cells with minimal RNA degradation. The first method then employs a standard phenol:chloroform extraction to purify the RNA, whereas the second method uses a silica gel membrane to capture the RNA.

Total RNA was isolated from 10^6 BHK cells following the 2 protocols. The resulting RNA samples were separated on a denaturing agarose gel for comparison and also analysed by spectrophotometry. This showed that both methods returned similar quality RNA but that phenol:chloroform isolation resulted in increased yields of RNA (data not shown). In addition, the yield of RNA isolated in this way was further improved when all vortexing steps were increased to 1min. Therefore, all subsequent phenol:chloroform extractions used increased vortex times.

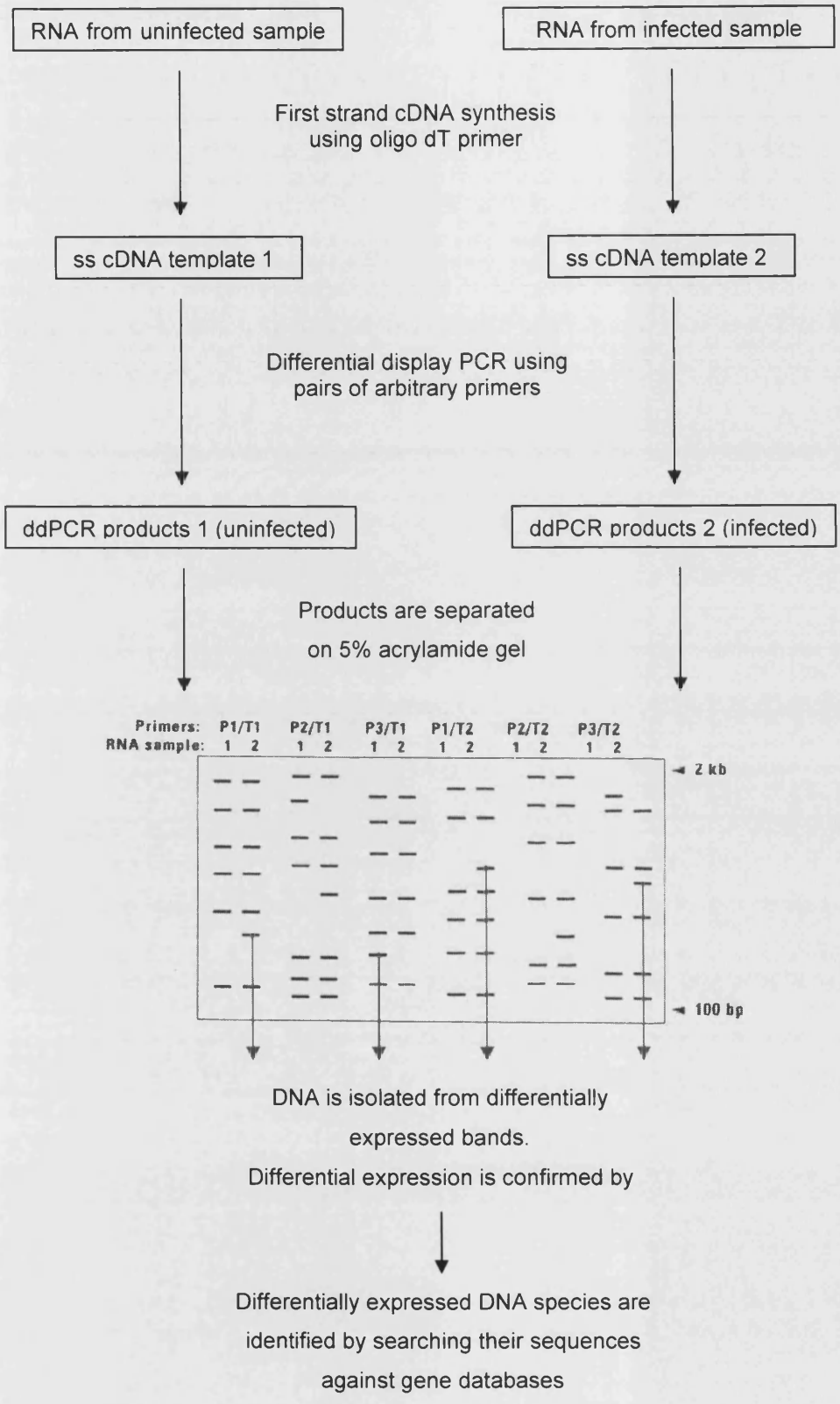


Figure 3.1. Overview of the differential display protocol for identification of differentially expressed genes during MHV-68 infection.

An example of RNA quality analysis, by separation on a denaturing agarose gel, is shown in Figure 3.2A. Both lanes 1 and 2 represent the same RNA sample, but at different quantities. Lane 2 was analysed by densitometry in Figure 3.2B. The 2 largest peaks represent the 28S (4.8kb) and 18S (1.8kb) ribosomal RNA (rRNA) subunits. The low molecular weight peak consists of the 5S rRNA subunit and tRNAs. Finally, the 2 small high molecular-weight peaks represent the 45S and 32S rRNA precursors. If all of these peaks are detected, then the RNA is likely to be of good quality. Ideally, each peak should be well defined and sharp, indicating minimal degradation of RNA. In addition, the 28S rRNA peak should be higher than the 18S rRNA peak.

Each RNA sample was also analysed by spectrophotometry. Absorbance at 260nm was used to calculate yield. Absorbance at 280nm was used to indicate protein contamination; samples showing A_{260}:A_{280} ratios below 2 were discarded as this indicated protein contamination.

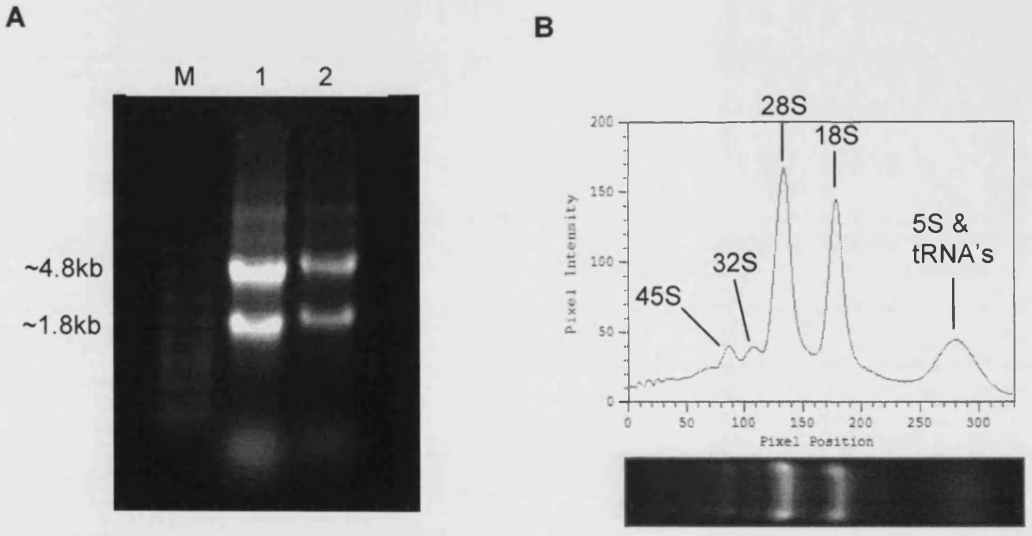


Figure 3.2. Denaturing agarose gel analysis of RNA samples.
A. Denaturing 1% agarose gel analysis of RNA isolated using TRIzol. RNA isolated from 5 x 10^5 BHK cells was electrophoresed in Lane 1. RNA isolated from 2.5 x 10^5 BHK cells was electrophoresed in Lane 2. M represents the RNA marker (Invitrogen). **B.** Densitometric analysis of lane 2 (UVIsoft gel analysis software, UVItech). The peaks on the graph represent the ribosomal subunits. The identity of each ribosomal subunit is indicated.

3.2.1 Minimising DNA Contamination

To minimise DNA contamination of RNA, all samples were incubated with DNAse I. PCR amplification of RNA template (Figure 3.3) was used to show that any DNA contamination had been removed. As the differential display was also PCR-based, this was a valid measure of DNA contamination for this system. Therefore, all subsequent RNA samples for the differential display were DNAse treated.

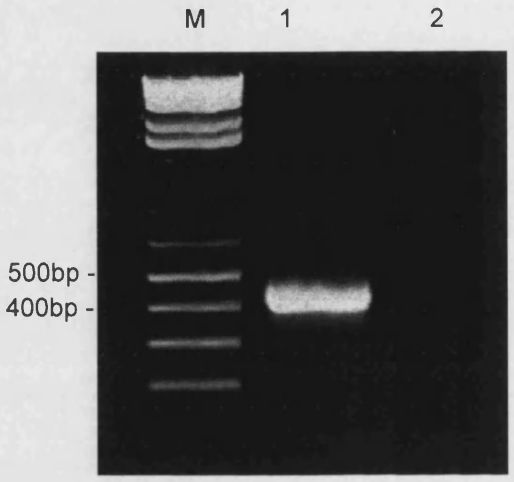


Figure 3.3. Assessing DNA contamination of RNA samples.
RNA was isolated from BHK cells and then incubated with DNAse I to remove any DNA contamination. PCR amplification, using primers against β-actin, showed that DNA contamination was removed by DNAse treatment. Lane 1 shows PCR amplification β-actin sequence from cDNA template, which was reverse transcribed from DNAse-treated RNA. Lane 2 shows PCR amplification β-actin sequence from DNAse-treated RNA template. M indicates the DNA marker (1kb+ marker, Invitrogen).

3.3 cDNA Synthesis

3.3.1 Yield of cDNA

The RNA samples are reverse transcribed to cDNA, which the differential display system uses as template to produce a "display". To increase the quality and quantity of the cDNA template, the incubation time for cDNA synthesis was increased from 15min to 1h, as suggested in some protocols (Ralph, 1993; Tal-

Singer, 1998). Extending the incubation time would also increase the probability of rare RNAs being reverse transcribed.

RNA was reverse transcribed for 15min and 1h and separated on an agarose gel to confirm that cDNA yield was increased when incubation time was increased. Lanes 1 and 2 of Figure 3.4 show cDNA produced following 15min incubation, while lanes 3 and 4 show the results of 1h incubation. As more cDNA was produced when 1h incubation was used, subsequent cDNA was synthesised with 1h incubation.

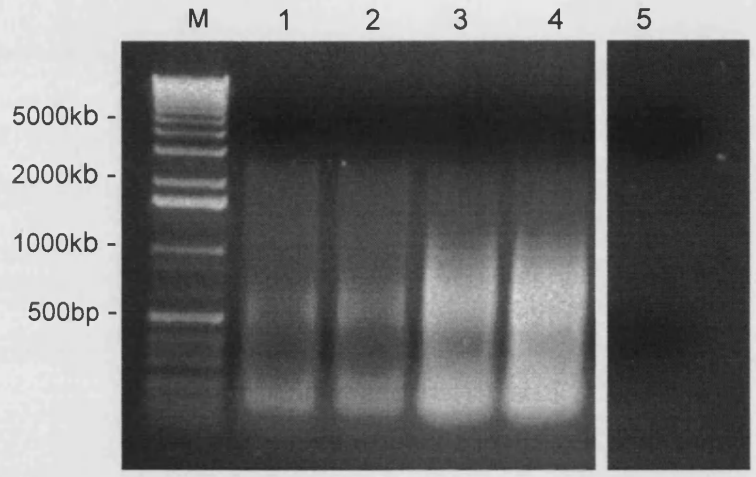


Figure 3.4. Assessing yield of the cDNA synthesis reaction.
RNA was reverse transcribed using two different incubation periods. Both reactions were primed with oligo-dT. Lanes 1 & 2 show the cDNA resulting from a reverse transcription with a 15min incubation. Lanes 3 & 4 show the cDNA resulting from a reverse transcription with a 1h incubation. Lane 5 shows a control where H_2O was used instead of RNA template. M indicates the DNA marker.

3.3.2 Integrity of cDNA

In Figure 3.4, single stranded DNA (ssDNA) was electrophoresed through an agarose gel. However, the DNA marker that was used was double stranded DNA (dsDNA). As these two forms of DNA possess different electrophoretic properties, it was not possible to be certain of the size range of cDNA produced by reverse transcription. As the size range of cDNA synthesised from RNA is an indicator of the quality, it is useful to ascertain the size of the cDNAs produced. Typically, cDNA produced from mammalian poly-A RNA ranges from 0.5 – 10kb in size (Chenchik,

1995). Therefore, a similar size range was expected for cDNA derived from BHK cell RNA.

The single stranded cDNA shown in Figure 3.4 was therefore used to synthesise double stranded cDNA. This was electrophoresed through an agarose gel as shown in lane 1 of Figure 3.5. As a positive control, double stranded cDNA was synthesised from human placental poly-A RNA (Clontech) and this is shown in lane 2. The dsDNA produced from BHK RNA ranged from around 100bp to more than 10kb. The control reaction with human placental poly-A RNA showed similar results. This suggested that RNA of all sizes was successfully reverse transcribed to cDNA.

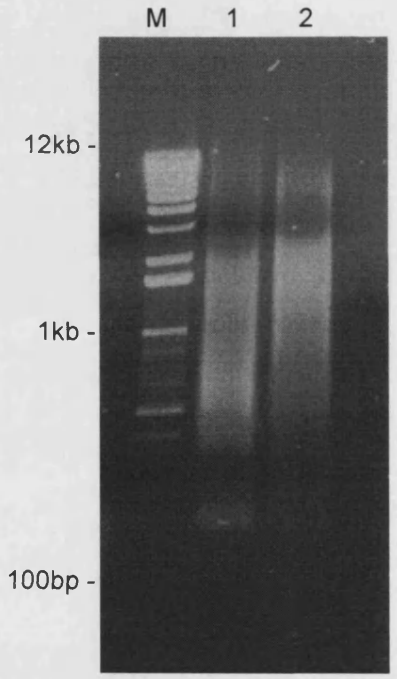


Figure 3.5. Size range of cDNA synthesised from BHK cell RNA.
dsDNA was synthesised from BHK cell RNA (lane 1), and as a control, from placental poly-A RNA (Clontech; lane 2). The products were separated on a 1.2% agarose gel. M indicates the DNA marker.

3.3.3 Identification of the Limiting Factor in cDNA Synthesis

The previous experiments confirmed that cDNA synthesis was efficient, However, it was also important that the only limiting factor in the reaction was the amount of RNA template. This would ensure that reverse transcription was not terminated prematurely due to a lack of reagents, and would result in a cDNA population that was representative of the original poly-A RNA population. Therefore, three-fold serial dilutions of RNA (equivalent to 1×10^5, 3.3×10^4 and 1×10^4 BHK cells, respectively) were used as template to produce cDNA.

The reaction products were electrophoresed on a 2% agarose gel, as shown in Figure 3.6A, lanes 4 - 6. The intensity of each cDNA smear was calculated by densitometry (UVGelWorks software) to provide an estimate of the yield of each reaction. Plotting these measurements on a graph, as shown in Figure 3.6B, showed a linear relationship between the amount of RNA template and yield of cDNA ($R^2 = 0.995$). This showed that RNA concentration was the limiting factor for cDNA synthesis, when the RNA isolated from $1 \times 10^4 - 1 \times 10^5$ BHK cells was used.

The reverse transcription reaction was performed with 15min (lanes 1 – 3) and 1h (lanes 4 – 6) incubations. This confirmed again that increasing incubation time had a significant effect on yield, as the cDNA smears seen in lanes 4 – 6 were more intense than those in lanes 1 – 3.

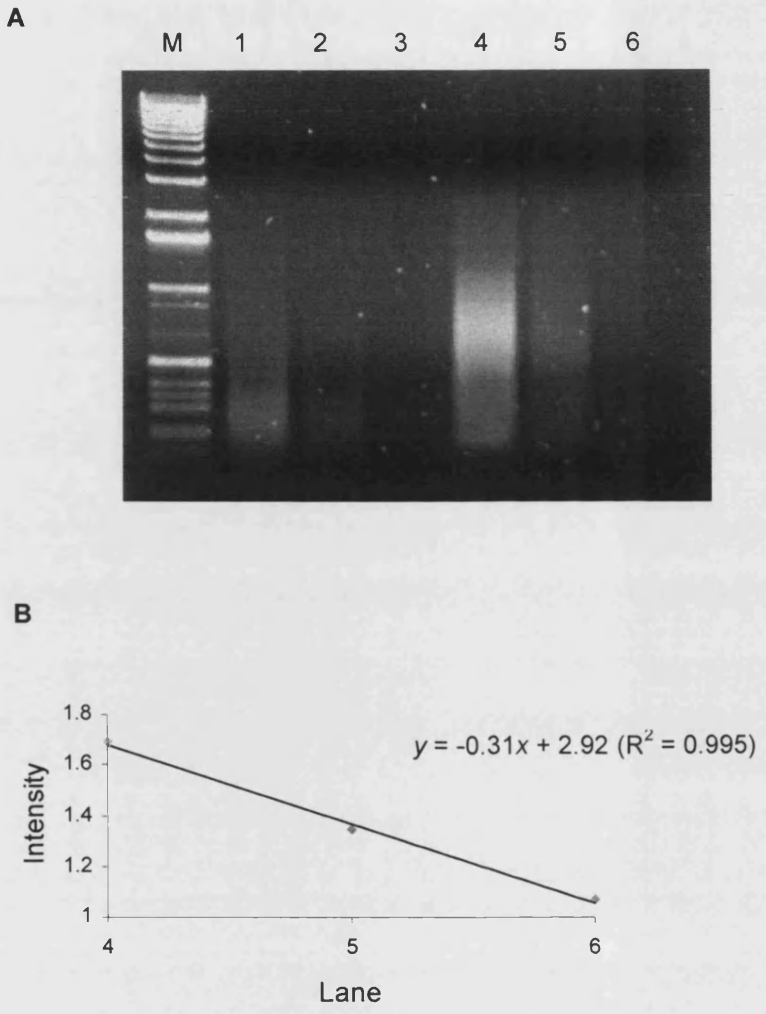


Figure 3.6. Effect of RNA template concentration and incubation time on yield of cDNA.
A. Two sets of cDNA synthesis reactions were performed with a 15min (lanes 1 – 3) and 1h (lanes 4 – 6) incubation period. Each set of reactions used 3-fold dilutions of RNA, corresponding to 1 x 10^5, 3.3 x 10^4 and 1 x 10^4 BHK cells, respectively. B. Linearity of cDNA synthesis. The cDNA yield was estimated by calculating signal intensities of each lane using UVGelWorks software. The results for lanes 4-6 were plotted to show that the relationship between quantity of RNA template and cDNA yield was linear.

3.4 Differential Display of MHV-68 Infected and Uninfected Cells

3.4.1 Trial Run

As isolation of RNA and synthesis of cDNA was now optimised, the differential display PCR (ddPCR) step was investigated. RNA was isolated from 10^5 BHK cells and reverse transcribed. The Delta differential display protocol (Clontech) was used to provide primer sequences and reaction conditions for the ddPCR. The cDNA was diluted 1:10 and 1:40, and each dilution was PCR amplified separately. This served as an internal control and only results duplicated in both PCR amplifications would be considered valid.

Initially, 12 combinations of 5 degenerate primers (primers P1, P2, P3 and T8, T9) were used to amplify the cDNAs. P-primers refer to general degenerate primers and T-primers refer to anchored primers designed to bind to the poly-T tail of cDNA species. Unlabelled nucleotides were used for this trial run. Table 3.1 shows a scheme of the ddPCR reactions set up for this trial run.

The ddPCR was performed and each ddPCR amplification product was electrophoresed on a 1% agarose gel, as shown in Figure 3.7. A key to the reaction conditions represented in each lane of the gel can be found in Table 3.2 Each primer pair was used to amplify the 2 dilutions of cDNA, and also a water template reaction to act as a negative control. Initial analysis of the gel suggested that different primer pairs produced different amplification products. Also many more bands were seen on the gel for some primer combinations than others. For example the primer combination of P3T9 gave many bands while the P2T8 primer pair gave few bands.

Densitometry was performed on the gel image and lane profiles were calculated. Sets of 3 lanes representing each primer pair with the 2 dilutions of cDNA, and the relevant negative control were plotted together as shown in Figure 3.8. These lane profiles showed that the same products resulted from both cDNA template concentrations, and therefore each band could be considered a valid result. The water negative control lanes showed that there was little background signal present.

Another group of controls used BHK RNA as template for the PCR amplification. Lanes 19 & 20 show these controls. A few bands can be seen in these lanes. Any bands of the same size were disregarded. However, overall there were very few bands in the RNA control lanes, which increased confidence in the system.

Table 3.1. ddPCR trial run setup.

Tube Label	cDNA Sample	Primers	Tube Label	cDNA Sample	Primers
Experimental displays			Water controls for each primer pair		
1	1A	P1 & T9	H1	Water	P1 & T9
2	1B	P1 & T9	H2	Water	P1 & T8
3	1A	P1 &T8	H3	Water	P2 & T9
4	1B	P1 &T8	H4	Water	P2 & T8
5	1A	P2 & T9	H5	Water	P3 &T9
6	1B	P2 & T9	H6	Water	P3 &T8
7	1A	P2 & T8			
8	1B	P2 & T8			
9	1A	P3 &T9			
10	1B	P3 &T9	Total RNA controls for each RNA		
11	1A	P3 &T8	R1	RNA	P1 & T9
12	1B	P3 &T8	R2	RNA/100	P1 & T9

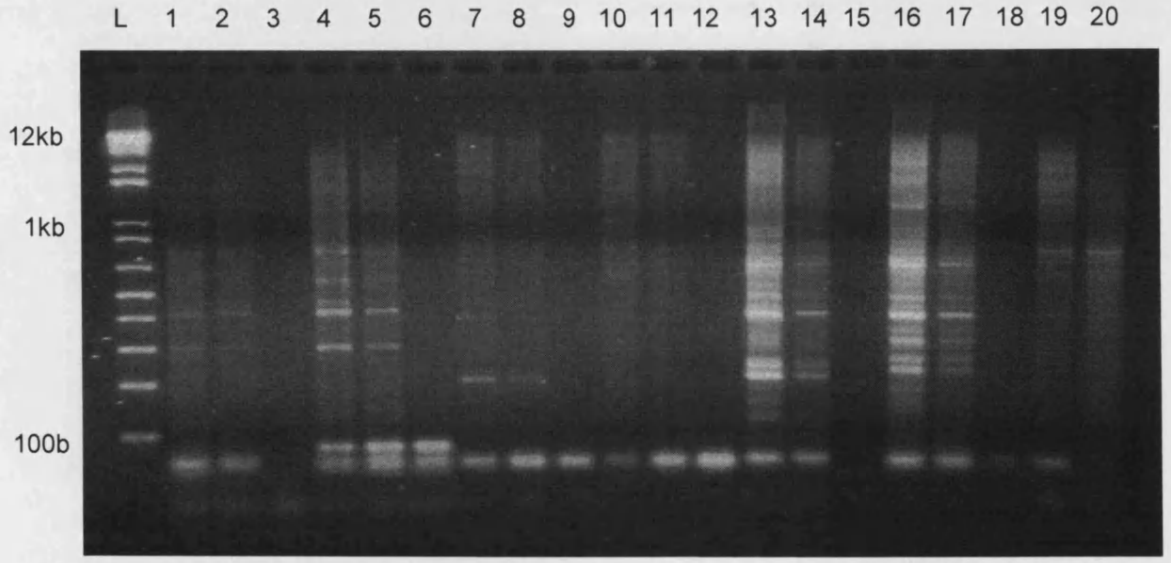


Figure 3.7. Results of the trial ddPCR run.

A key to the reactions shown in each lane can be found in Table **3.2**.

Table 3.2. Key to the ddPCR reactions shown in Figure 3.7.

Lane	1	2	3	4	5	6	7	8	9	10
Primers	**P1T9**			**P1T8**			**P2T9**			**P2T8**
Template	cDNA A	cDNA B	H2O	cDNA A	cDNA B	H2O	cDNA A	cDNA B	H2O	cDNA A

Lane	11	12	13	14	15	16	17	18	19	20
Primers	**P2T8**		**P3T9**			**P3T8**			**P1T9**	
Template	cDNA B	H2O	cDNA A	cDNA B	H2O	cDNA A	cDNA B	H2O	1ml RNA	0.01ml RNA

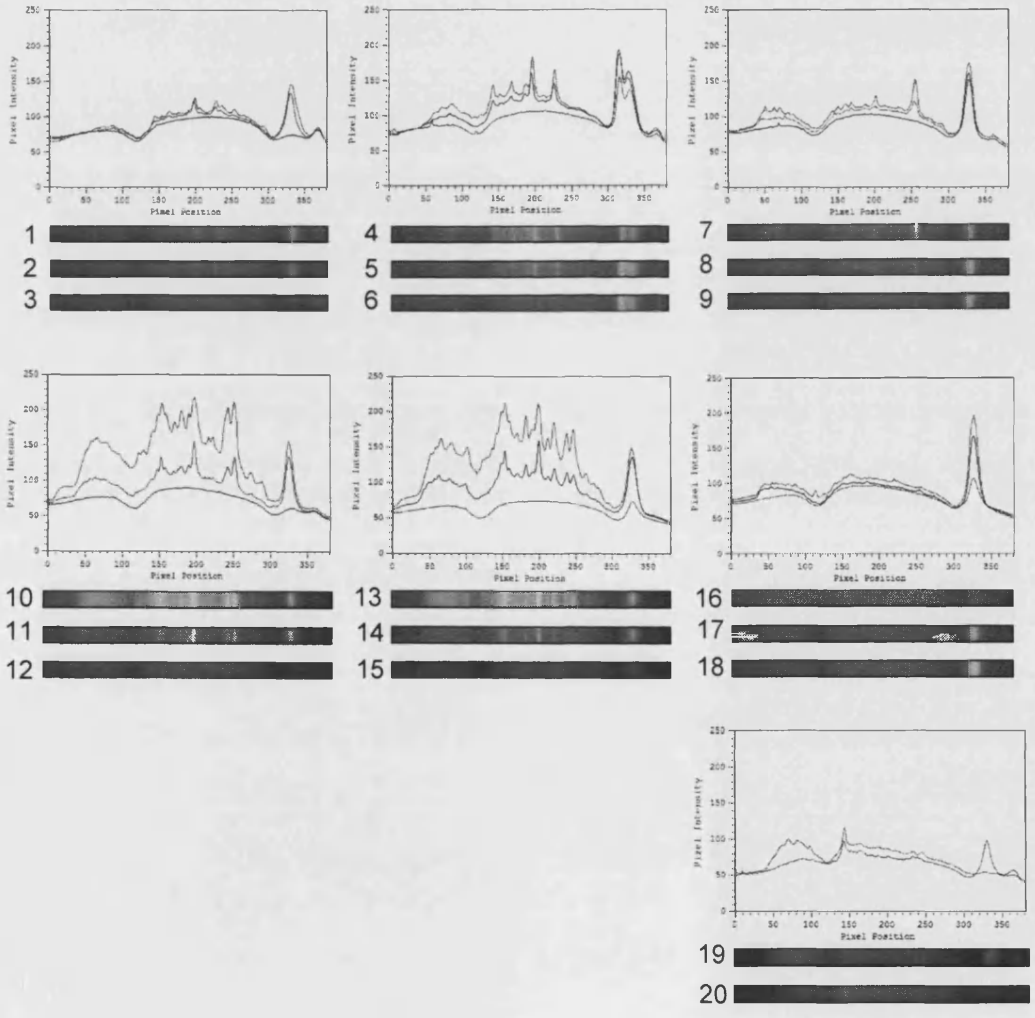


Figure 3.8. Lane profiles of trial ddPCR. Lane intensities were plotted to show the reaction products clearly, as well as the effects of diluting template.

3.4.2 Radiolabelled Trial Run

As the ddPCR was producing a display, a second trial run was performed using radiolabelled nucleotides. For this trial, RNA was isolated from BHK cells 24h pi with MHV-68 (moi 10). As RNA was being isolated towards the end of the MHV-68 life-cycle, the accumulation of a significant quantity of viral RNA was expected. Mock-infected BHK cells were used as a negative control. The experimental design was identical to that used for the first trial run, except for the inclusion of [α-^{33}P] dATP in the reaction mix. The reaction details are summarised in Table 3.3.

The ddPCR reaction products were separated on a sequencing gel to allow greater resolution of the display. The gel was dried and used to expose both film and a phosphor screen. The phosphoimager was found to give a clearer and sharper picture than the film. The image, shown in Figure 3.9, shows 6 groups of 5 lanes each, which consist of 2 uninfected lanes (both cDNA concentrations), 2 infected lanes (both cDNA concentrations) and a negative control lane (water template), respectively. The final group of 6 lanes represent the RNA controls, one for each primer combination used.

There were many examples of differentially displayed bands in each group of lanes. Some of these are highlighted with black boxes in Figure 3.9. These bands, which were present in one lane and absent in another, are examples of those that could be analysed further, to potentially identify differentially expressed genes.

Again the two dilutions of each cDNA template gave the same profile, as seen in Figure 3.10A and Figure 3.10B, which shows the lane profiles of small sections of the gel. Differences between infected and control cells can be seen in panels C & D of Figure 3.10.

Unfortunately, only bands in the top half of the gel were clearly visible. To overcome this problem, each ddPCR reaction product was electrophoresed twice, once for 2h and once for 4h, to give a short and long run. The two runs allowed larger and smaller amplification products, respectively, to be analysed at the same time.

Table 3.3. ddPCR hot run setup.

Lane	1	2	3	4	5	6	7	8	9	10
Primers	**P1T9**					**P1T8**				
Template	Uninf.* A*	Uninf. B*	Inf.* A	Inf. B	H_2O	Uninf. A	Uninf. B	Inf. A	Inf. B	H_2O

Lane	11	12	13	14	15	16	17	18	19	20
Primers	**P2T9**					**P2T8**				
Template	Uninf. A	Uninf. B	Inf. A	Inf. B	H_2O	Uninf. A	Uninf. B	Inf. A	Inf. B	H_2O

Lane	21	22	23	24	25	26	27	28	29	30
Primers	**P3T9**					**P3T8**				
Template	Uninf. A	Uninf. B	Inf. A	Inf. B	H_2O	Uninf. A	Uninf. B	Inf. A	Inf. B	H_2O

Lane	31	32	33	34	35	36
Primers	**P1T9**	**P1T8**	**P2T9**	**P1T9**	**P1T8**	**P2T9**
Template	RNA A	RNA A	RNA A	RNA B	RNA B	RNA B

Uninf* – uninfected

Inf* – infected

A* - less diluted 1:10 cDNA samples

B* - more diluted 1:40 cDNA samples

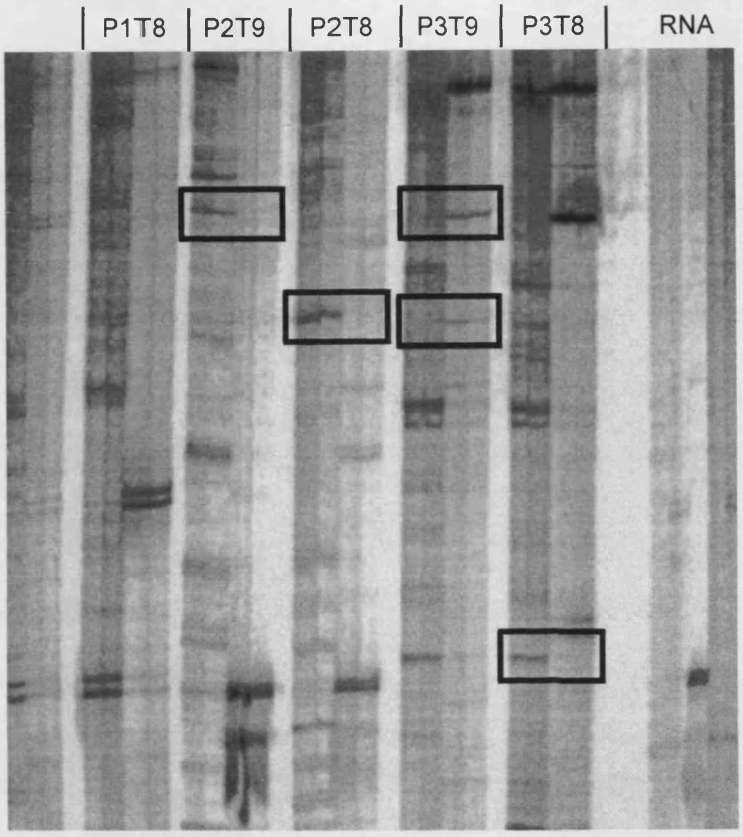


Figure 3.9. Differential display of MHV-68 infection.

BHK cells were infected for 24h at moi 10. RNA was isolated from these cells, as well as from a mock-infected control sample, and reverse transcribed. Each resulting cDNA sampled was diluted to 1:10 (A) and 1:40 (B) and amplified by ddPCR. The reaction products were electrophoresed for 2h and used to expose a phosphor screen. A section of the resulting image is shown. The primers used in each group of reactions are shown above each panel. The details of each reaction are summarised in Table 3.3. Each of the 6 groups of 5 lanes consists of uninfected cDNA dilution A, uninfected cDNA dilution B, infected cDNA dilution A, infected cDNA dilution B and a water control, respectively. The last group of 6 lanes consists of the RNA controls, with one lane for each primer pair. Examples of differentially expressed bands have been highlighted with black boxes.

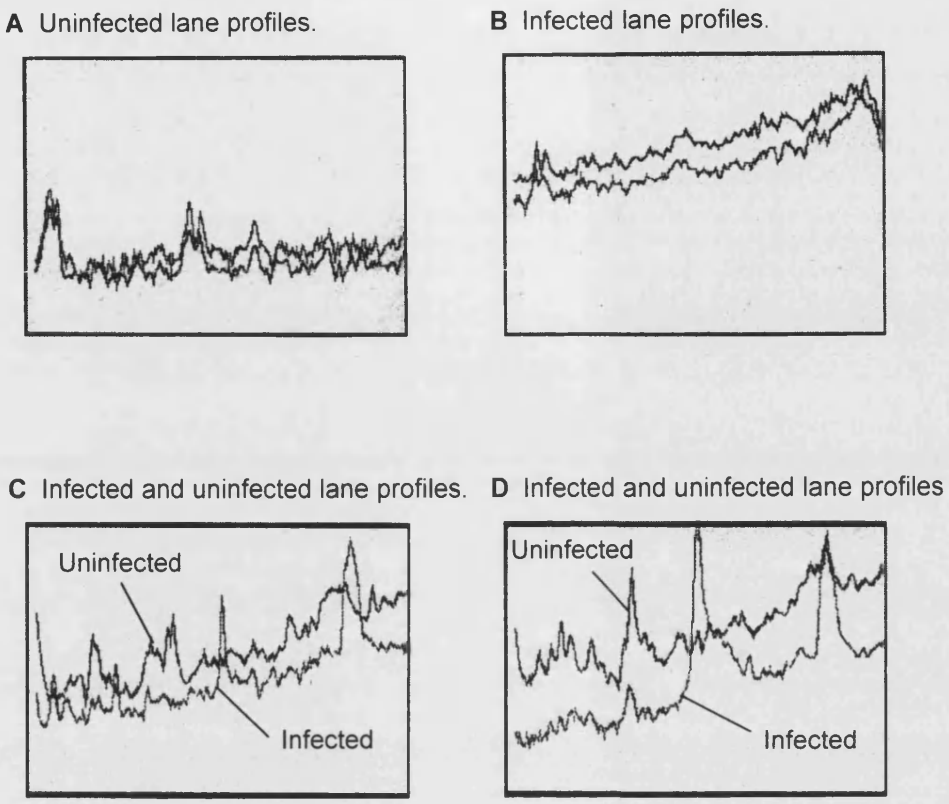


Figure 3.10. Lane profiles of radiolabelled ddPCR trial run.
Sections of the gel were analysed using Bas software (Fuji). The profiles of 2 lanes are shown in each panel. A. Profiles of two adjacent lanes, both uninfected, showing the amplification of 2 concentrations of the same cDNA. B. Profiles of two adjacent lanes, both infected, showing the amplification of 2 concentrations of the same cDNA. C. & D. Contrasting profiles of an uninfected and infected lane, showing differentially expressed RNAs.

3.4.3 Experimental Run

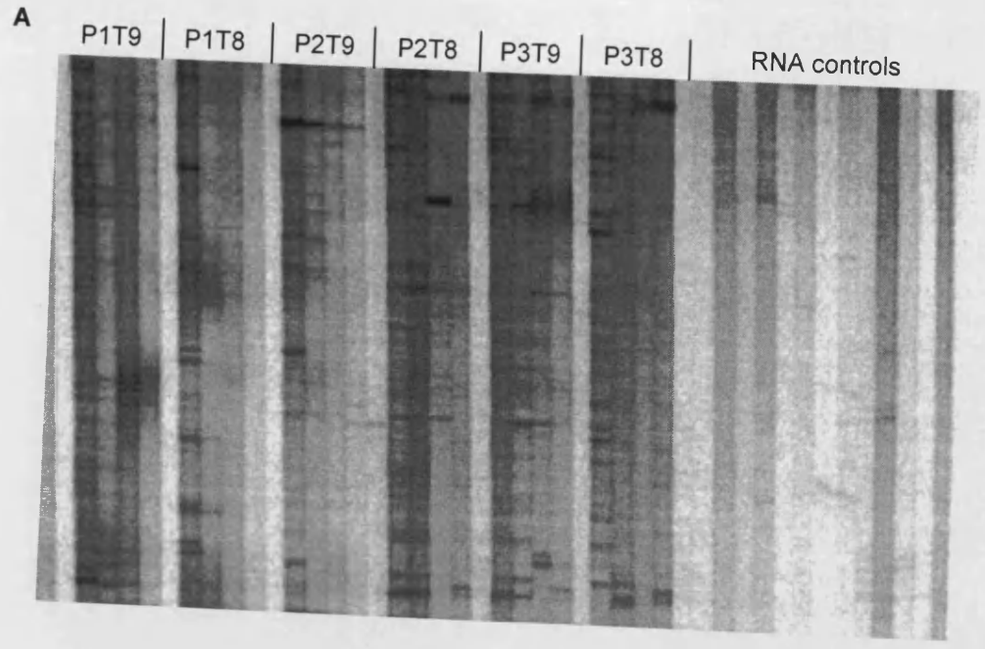
RNA was isolated from MHV-68 infected (moi 10) BHK cells and mock-infected BHK cells 24h pi, as described previously. cDNA was synthesised and used as template at two dilutions for the ddPCR amplification. The reaction products were separated on a sequencing gel for both 2h (short run) and 4h (long run). The gel was then dried and visualised by exposure to both film and a phosphor screen.

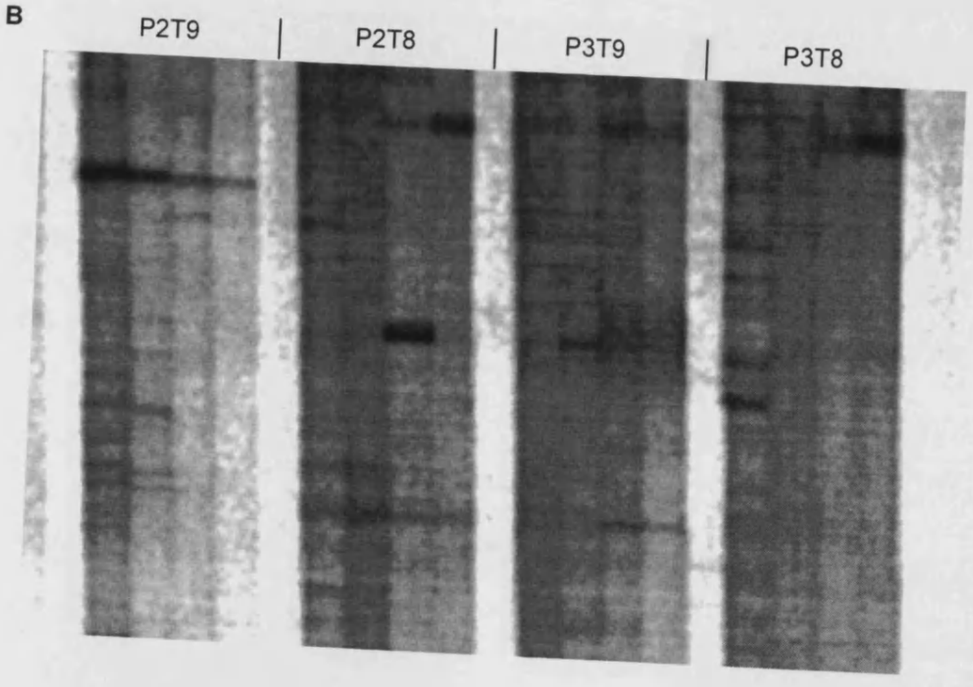
A scan of the gel is shown in Figure 3.11. As before there are 6 groups of 5 lanes followed this time by a full set of 12 RNA control lanes. There were a number of global differences between infected and uninfected lanes. The uninfected lanes had

more bands than the infected ones, suggesting that there was a greater diversity of host-derived RNAs than viral ones. Most bands present in the uninfected lanes were absent in the infected lanes, suggesting that host protein synthesis shutoff was occurring. There were a significant number of bands present only in the infected lanes, which indicated that many viral RNAs were present in infected cells 24h pi. There were also some bands that appeared to be present in both the uninfected and infected lanes, mostly at a lower intensity in the infected lanes. This suggested that some host-derived mRNAs were either still being transcribed at a lower rate than in uninfected cells, or that their transcripts had not degraded 24h pi.

Initially, 11 bands were selected for further analysis. Eight of these are shown in Figure 3.12. These differentially expressed bands fell into the following categories: Bands present only in uninfected lanes (Figure 3.12A & Figure 3.12B). Bands present only in infected lanes (Figure 3.12C & Figure 3.12D). Bands present in both uninfected and infected lanes but of reduced intensity in the latter (Figure 3.12E & Figure 3.12F). Bands present in both uninfected and infected lanes but of increased intensity in the latter (Figure 3.12G & Figure 3.12H).

Figure 3.11 (following page). Differential display of MHV-68 infection.
BHK cells were infected for 24h at moi 10. RNA was isolated from these cells, as well as from a mock-infected control sample, and reverse transcribed. Each resulting cDNA sampled was diluted to 1:10 and 1:40, and amplified by ddPCR. The reaction products were electrophoresed for 2h and 4h, to aid separation of the large range of product sizes. The gel was dried and visualised on a phosphoimager. A. A section of the resulting image is shown. B. Part of the image in panel A is shown magnified. The primers used in each group of reactions are shown above each panel. Each of groups of lanes consists of 2 lanes corresponding to differential display amplification of uninfected samples (using 2 dilutions of cDNA dilution 1:10 and 1:40), 2 lanes corresponding to differential display amplification of infected samples (using 2 dilutions of cDNA dilution 1:10 and 1:40), and the fifth lane show a water template negative control. The last group of 12 lanes consists of the RNA controls, with one lane for each primer pair.


A
P1T9
P1T8
P2T9
P2T8
P3T9
P3T8
RNA controls


B
P2T9
P2T8
P3T9
P3T8

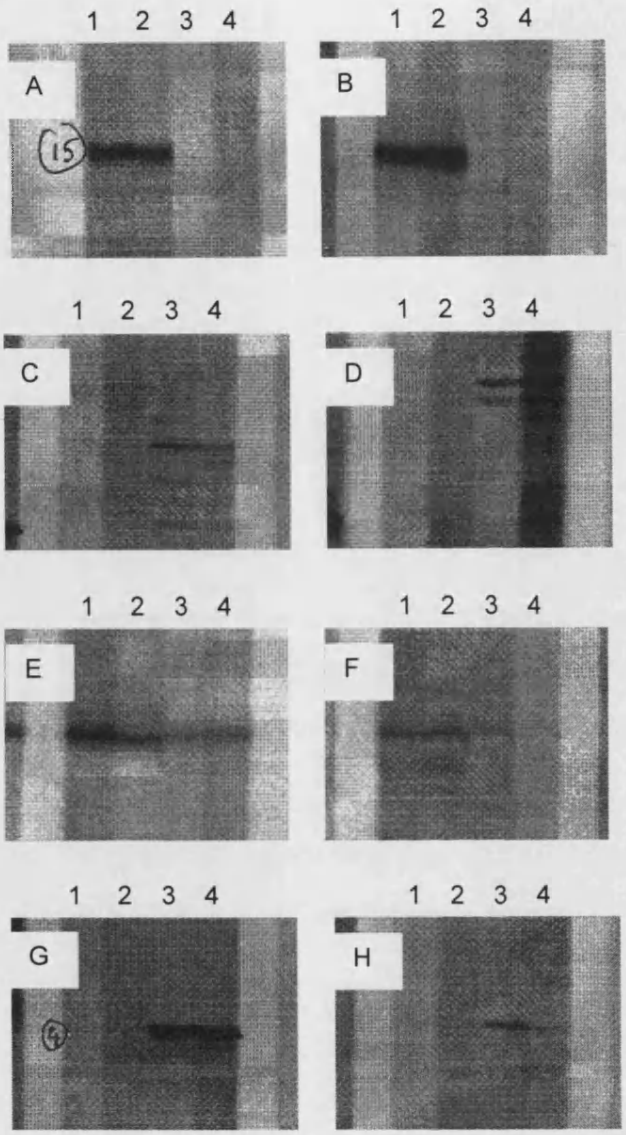


Figure 3.12. Examples of differentially expressed bands.

Each panel shows an example of a band that was differentially amplified from mock-infected and infected samples, as shown in Figure 3.11. Each panel shows 4 lanes representing uninfected cDNA dilution A (lane 1), uninfected cDNA dilution B (lane 2), infected cDNA dilution A (lane 3), infected cDNA dilution B (lane 4). A fifth lane (not marked) shows the water template control. A. & B. Bands representative of genes only expressed in mock-infected cells. C. & D. Bands representative of genes only expressed in infected cells. E. & F. Bands representative of genes expressed in mock-infected cells, but also to a lesser degree in infected cells. G. & H. Bands representative of genes expressed in mock-infected cells, but expressed at higher levels in infected cells.

3.5 Elution of DNA from Differentially Displayed Bands

The bands selected for further analysis were marked on the autoradiography film and then lined up carefully with the gel. The bands were excised from the gel and the DNA eluted by boiling the gel fragments in TE buffer. The eluted DNA was reamplified by PCR reaction using the appropriate primers, which had produced the corresponding band at the ddPCR stage. To check that the bands had been correctly excised, the gel was used to expose a phosphor screen a second time and the absence of the selected band confirmed.

DNA was initially eluted from gel fragments by boiling in TE buffer for 20min. This DNA solution was used to reamplify the bands of interest. The reamplification products are shown in Figure 3.13. Only 6 of 11 bands were successfully reamplified. As an attempt to optimise the elution protocol, the incubation period for gel fragments in TE buffer was reduced to 10min. This resulted in an increased success rate for reamplification of bands. Of 14 bands excised from the polyacrylamide gel, 13 were successfully reamplified. The reamplification products are shown in Figure 3.14. Further DNA elutions were performed with a 10min incubation.

In some instances, several bands resulted following the reamplification of the DNA eluted from a single original band, as shown in Figure 3.14. This was due to the degenerate nature of the differential display primers used for this reamplification: The degenerate primers were designed to bind to more than one site on the DNA template. Also, more than one DNA species could be observed as a single band on the polyacrylamide gel if they possessed similar electrophoretic properties. In fact, both these phenomenon could be occurring at the same time.

If more than one DNA species was present in a single original band on the polyacrylamide gel, it was not clear which DNA species had been differentially expressed. Therefore, it was necessary to take all the reamplification products for further analysis and confirm which of these represented a differentially expressed gene. This was achieved by northern blot analysis.

Figure 3.15 shows the results of northern blot analysis for the reamplified DNA shown in lane 7 of Figure 3.14. A single DNA species was gel purified from the band seen in Figure 3.14, and used to produce a radiolabelled probe. The northern blot analysis confirmed that the DNA species represented a differentially expressed

gene that was only expressed in MHV-68 infected cells. While most probably a viral gene, cloning and sequencing were required to positively identify this gene.

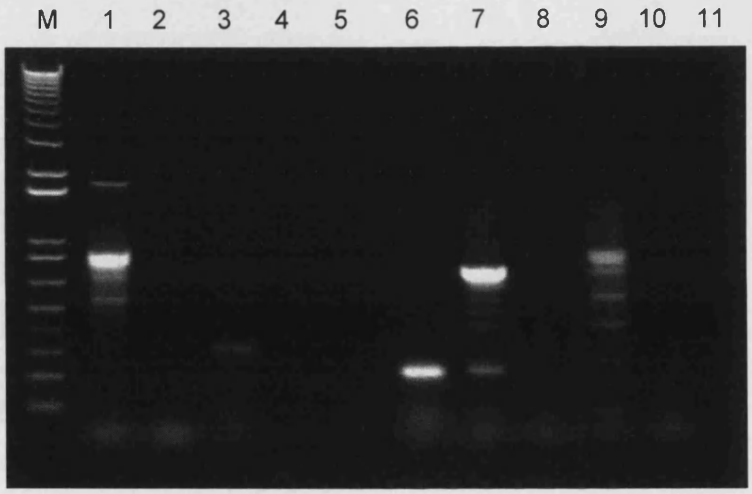


Figure 3.13. Reamplification of differentially expressed bands.
Bands representing differentially expressed genes were excised from dried polyacrylamide gels. The DNA was eluted from these gel fragments by boiling for 20min in TE buffer. The eluted DNA was then reamplified by PCR using the corresponding primers that had produced the band at the ddPCR stage (see Table 3.3). Lanes 1 – 11 represent 11 bands that were selected for further analysis. M indicates the DNA marker.

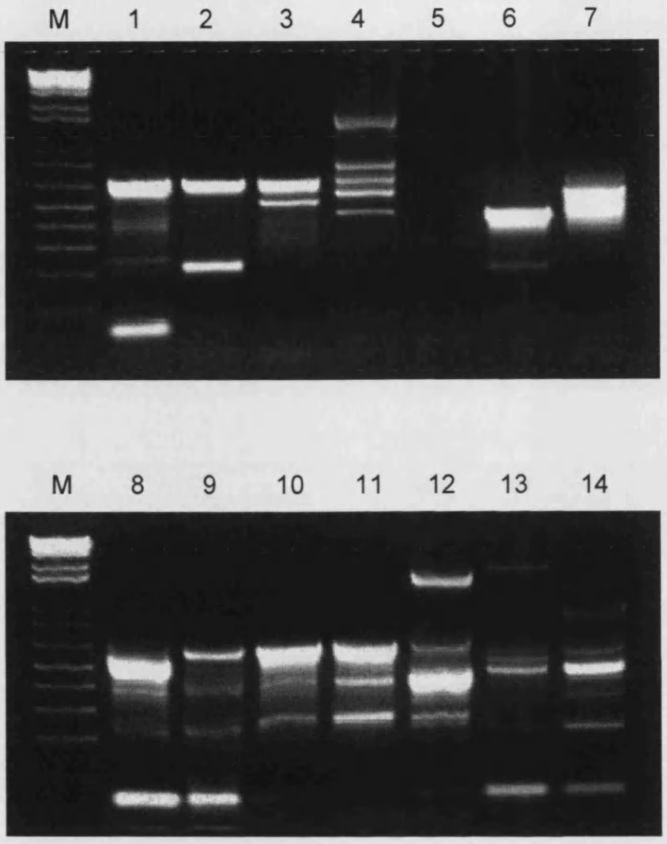


Figure 3.14. Further reamplification of differentially expressed bands.
Bands representing differentially expressed genes were excised from dried polyacrylamide gels. The DNA was eluted from these gel fragments by boiling for 10min in TE buffer. The eluted DNA was then reamplified by PCR using the corresponding primers that had produced the band at the ddPCR stage (see Table 3.3). Lanes 1 – 14 represent 14 bands that were selected for further analysis. M indicates the DNA marker.

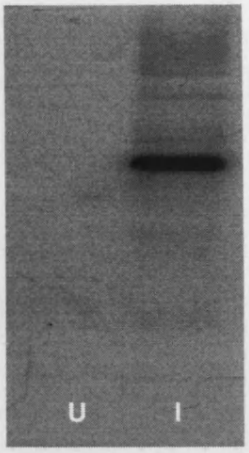


Figure 3.15. Northern blot analysis to confirm a differentially expressed gene identified by ddPCR.

The DNA from band showing differentially expression was eluted from the ddPCR polyacrylamide gel and reamplified. The DNA shown in lane 7 of Figure 3.14 was used to probe a northern blot of RNA from cells infected with MHV-68 (moi 10, 24h pi; marked as I) and mock-infected cells (marked as U).

3.6 Sequencing of Differentially Displayed DNA

For practical reasons, potential differentially expressed genes were cloned and sequenced, and then if they could be identified, were used to probe northern blots. Therefore, the bands seen in Figure 3.14 were gel purified and ligated into the pGEM-T Easy vector system for cloning. As this cloning system uses a plasmid with a multiple cloning site within a β-galactosidase cassette, blue/white screening was used to select 2 transformed bacterial colonies for each cloned DNA species. Colony PCR was used to further confirm that cloning had been successful.

Initially 8 bands were cloned. The results of the colony PCR for 16 bacterial colonies are shown in Figure 3.16. This showed that 13 out of 16 colonies contained an insert. Of the remaining 3, lane 3 appeared to show a mixed colony, as 2 bands were present in this lane. Lanes 14 & 16 had bands of approximately 170bp, which indicated that no insert was present (175bp = the distance between the primers binding to a pGEM vector without an insert). Therefore, the culture represented by lane 3 was streaked out and new colonies picked. Colony PCR was

used again to confirm the presence of an insert. The cultures represented by lanes 14 & 16 were discarded.

All cultures containing successfully cloned DNA species were grown up and plasmid extracted. The purified plasmid was then used as template for sequencing. The resulting sequence data was used to perform BLAST searches to identify the potential differentially expressed genes.

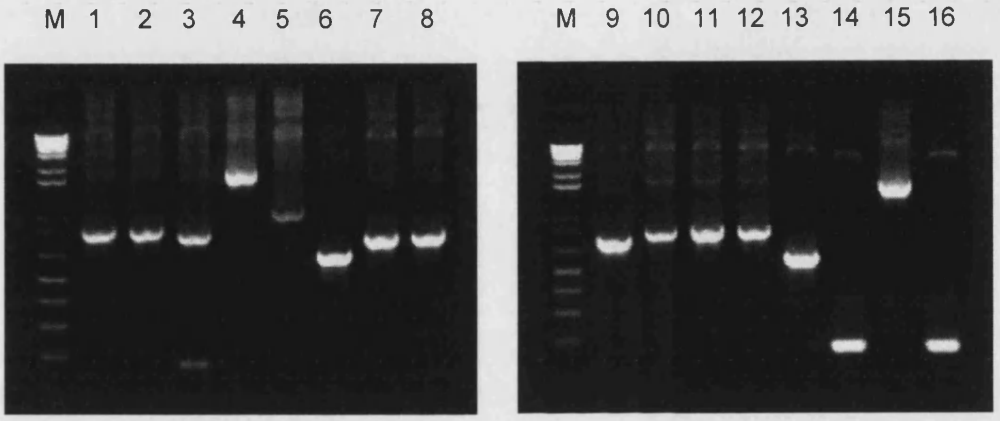


Figure 3.16. Colony PCR to check for presence of an insert.
Blue/white screening was used to select colonies containing insert-containing plasmids. These colonies were then used as template for colony PCR to confirm the blue/white screening. SP6 and T7 primers (see Appendix I) that bound to either side of pGEM multiple cloning site were used. Absence of an insert was indicated by a band of 175bp, which corresponded to the distance between the primer binding sites for plasmids without an insert. Lanes 1 – 16 represent white colonies of transformed bacteria.

3.7 Querying Public Databases via BLAST

For all DNA species that were successfully sequenced, the data was used to query the National Centre for Biotechnology Information (NCBI) databases via BLAST. Sequences that had matches in the NCBI databases were used to probe northern blots to confirm their differential expression. The results of these database searches and subsequent northern blot analyses are summarised in Table 3.4. Examples of these northern blot analyses are shown in Figure 3.17.

Of the potential differentially expressed genes that were successfully identified through BLAST, all genes whose expression had increased 24h pi were identified

as viral, e.g. MHV-68 ORF 25, 57 and 66. All the genes whose expression was decreased following infection were found to be host genes, e.g. adenine phosphoribosyltransferase and fibronectin. A degree of redundancy appeared in the results, as a number of potential differentially expressed genes were found to be the same gene. For example, MHV-68 ORF 25 (major capsid protein) was identified 4 times from 3 independently isolated differentially displayed bands.

Table 3.4. Results of BLAST searches with sequenced, differentially expressed DNA species.

Band	BLAST result	Score[1]	P value[2]	Confirmation by northern
1	U97553: MHV-68 ORF25	1298	8.60E-100	Up
3	X03603: Mus pahari adenine phosphoribosyltransferase	528	8.10E-62	Down
4	X93167: Mus musculus mRNA for fibronectin	1134	3.20E-86	Down
7	U97553: MHV-68 ORF57	1045	8.70E-80	Up
8	U97553: MHV-68 ORF57	1021	8.50E-80	Up
10	U97553: MHV-68 ORF25	1008	3.70E-111	Up
12	AF111102: Mus musculus major histocompatibility complex class	129	4.00E-28	Down
14	AI449820: Stratagene mouse testis, Mus musculus cDNA clone 917572, mRNA sequence	103	5.60E-24	Down
17	U97553: MHV-68 ORF25	884	0	Up
19	U97553: MHV-68 ORF58	656	0	Up
20	U97553: MHV-68 ORF25	640	0	Up
24	NM_145507: Mus musculus similar to Aspartyl-tRNA Synthetase	54	6.00E-05	Down
25	U97553: MHV-68 ORF58 & 59	485	7.20E-92	Up
28	U97553: MHV-68 ORF66	264	6.00E-70	Up

[1]The score of an alignment, calculated as the sum of substitution and gap scores. Substitution scores are given by a look-up table and gap scores are typically calculated as the sum of the gap opening penalty and the gap extension penalty.

[2]The probability of an alignment occurring with the score in question or better. The p value is calculated by relating the observed alignment score to the expected distribution of high scoring pair scores from comparisons of random sequences of the same length and composition as the query to the database. The most highly significant P values will be those close to 0.

Band 25 listed in Table 3.4 is of particular interest as when submitted via BLAST to the databases, the nucleotide sequence was found to match both ORF 58 and 59 of MHV-68. This suggested that the DNA species in band 25 was synthesised from a RNA species that spanned the reading frames of ORF 58 and 59. It therefore seems likely that these two ORFs are transcribed together as part of a polycistronic mRNA.

There were a large number of sequences that did not have matches in the NCBI databases. Since the time of these searches (2000), the databases have grown as more genes have been sequenced. Therefore, the unresolved sequences, and also the previously identified genes, were resubmitted via BLAST to query the databases again.

The results are presented in Table 3.5. This confirmed that the databases have indeed become more comprehensive, reflecting in particular the concerted efforts of the mouse genome project. Several of the previously unidentified sequences now returned hits from the database, although there were still some sequences with no matches. Furthermore, bands 12 & 14 that had been identified in 2000 but with weak scores, now returned matches with higher scores. This illustrates the need to search the databases at regular intervals, until definitive identification is possible.

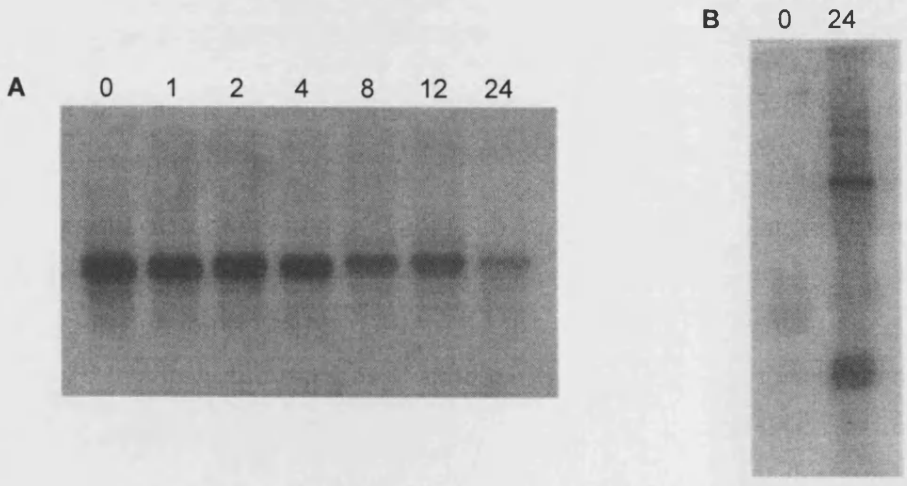


Figure 3.17. Northern blot confirmation of differential display.
Bands were excised from the ddPCR gel, amplified and sequenced, before being used to probe RNA blots. Band 14 is shown in A and band 1 (MHV-68 major capsid protein) in B, hybridised to blots with RNA harvested at various h pi indicated above each lane.

Table 3.5. Results of repeated BLAST searches following development of gene databases.

Band	BLAST result	Score[1]	P value[2]
11	AF109905.1: Mus musculus major histocompatibility locus class III regions Hsc70t gene, partial cds; smRNP, G7A, NG23, MutS homolog, CLCP, NG24, NG25, and NG26 genes, complete cds; and unknown genes	78	5.00E-12
12	Mus musculus DNA sequence from clone RP23-61O3 on chromosome 2, complete sequence	145	3.00E-32
13	No match found		
14	Mus musculus RIKEN cDNA 1700020L11 gene (1700020L11Rik), mRNA	410	1.00E-111
15	No match found		
16	Y08839.1: Mus auratus mRNA for delta-sarcoglycan, alternative first exon	123	9.00E-26
18	AF283763.1: Mus musculus thyrotropin-releasing hormone receptor 2 (TRH-R2) gene	66	3.00E-08
21	Mus musculus DNA sequence from clone RP23-338K8 on chromosome 11, complete sequence	42	0.26
22	Mus musculus DNA sequence from clone RP23-297J14 on chromosome 11, complete sequence	62	4.00E-07
23	No match found		
26	No match found		
27	Mus musculus clone RP24-491P21, complete sequence	38	7.1
29	No match found		

[1]The score of an alignment, calculated as the sum of substitution and gap scores. Substitution scores are given by a look-up table and gap scores are typically calculated as the sum of the gap opening penalty and the gap extension penalty.

[2]The probability of an alignment occurring with the score in question or better. The p value is calculated by relating the observed alignment score to the expected distribution of high scoring pair scores from comparisons of random sequences of the same length and composition as the query to the database. The most highly significant P values will be those close to 0.

3.8 Discussion

The differential display system was developed as a means of analysing changes in gene expression during MHV-68 infection *in vitro*. To assess the potential of this system for the study of MHV-68, the system was setup and optimised using a single time point pi. The resulting data was then analysed and used as an indication of the system's capabilities.

Each step of the differential display protocol (see Figure 3.1) was examined and optimised. RNA samples were treated with DNAse I and then thoroughly checked by denaturing gel electrophoresis and spectrophotometry to ensure that only high quality RNA template was used. Reverse transcription of RNA was optimised to increase the quality and yield of cDNA, whilst ensuring that the amount of template RNA was the limiting factor for this reaction.

Twelve primer pairs were used to assess the differential display system. The displays of RNA from MHV-68 infected BHK cells (24h pi) were compared against those from mock-infected controls. This showed clearly that the system was working and producing data. The results were further analysed to assess the potential of the system for investigating transcriptional events through MHV-68 infection.

Bands on the differential display gel could be categorised into those appearing only in lanes for infected samples, only in lanes for mock-infected samples, in both lanes but brighter in uninfected lanes, and in both lanes but brighter in infected lanes. The first group consisted of viral genes that were expressed by MHV-68 during infection. Several examples were isolated, sequenced and confirmed to be viral genes such as ORFs 25, 57 and 66, predicted to encode major capsid protein, an IE protein and capsid protein, respectively (see Table 3.4 and Table 3.5). The second and third groups consisted of host genes that were either no longer expressed or present at reduced transcript abundances 24h pi. This was presumably due to virally-induced shutoff of host protein synthesis. These 2 groups were not analysed further as these genes were the least interesting at this stage of the study. Of course identifying host genes that are rapidly shutoff at very early stages of infection could provide valuable data for further analysis.

Of those bands that appeared in both infected and uninfected lanes, those that showed increased abundance following infection were particularly interesting. Such an increased expression would suggest that these host genes in question were directly involved in the viral infection, either initiated by the host as a defence mechanism, or conversely by the virus to aid infection. However, it was not possible to confirm this differential expression via northern blot analysis. While the northern blot analyses showed that DNA species identified as viral genes were indeed all up-regulated following infection in agreement with the differential display gel, those DNA species identified as host genes were all found to be down-regulated following infection, although the differential display gels may have suggested otherwise.

This discrepancy could have arisen in one of three ways: Although the band in the infected and uninfected lane appeared to be the same DNA species, it is possible that they were in fact different DNAs of about the same size. As the RNA populations of infected and uninfected cells are very different, and as degenerate primers are used to initiate amplification, this is easily conceivable. It is also possible that more than a single DNA species may have been present in the gel fragment that was excised. Obviously this is undesirable and attempts were made to reduce the chances of this occurring by using a sensitive DNA labelling system and performing a long and short run for each ddPCR amplification product. Finally, several DNA species are reamplified from each gel fragment and only some of these may represent the original differentially expressed gene observed on the gel. Therefore, "non-signal" reamplified DNA species would be sequenced and used to probe northern blots. It became clear that some or all these potential pitfalls were occurring and therefore significantly increasing the noise in the differential display data.

These factors are potentially a major limitation of the differential display system. In particular the reamplification of DNA eluted from gel fragments was particularly problematic as it appeared to be difficult to cleanly isolate the DNA that was shown to be differentially expressed on the gel. Although the ddPCR successfully amplified the RNA samples, the polyacrylamide gel electrophoresis of the amplified products was not sufficient to separate them. The same type of gel electrophoresis is used successfully to separate sequencing reaction products, which only differ by 1 bp. Therefore, the differential display gel should also be capable of resolving 1bp differences. However, the ddPCR products are labelled with much higher levels of radioactivity and comprise a much larger size range. Optimisation of the electrophoresis could possibly further increase the resolution of the gel. However, this still leaves the other problems that increase the level of noise in the data. This noise results in an exponential increase in the workload required to isolate true signals from the differential display. It would be interesting to perform differential displays of infection at very early time points to see if these problems are as considerable when the 2 sample RNA populations are very similar.

There are also a number of limitations to the techniques used to interpret and analyse the differential display data: 0f 30 DNA species that were isolated from the differential display gel, only 23 were successfully sequenced (77% sequencing success). Of these, only 14 found matches in the databases (47% database matches). However, these limitations would be reduced with further optimisation of

the sequencing techniques and the continuing growth of the gene databases, as new genes are characterised. Indeed, of 13 resubmissions, 8 returned matches at this second attempt, which show a promising 62% success upon resubmission after a period of 2 years (listed in Table 3.5).

One unexpected result was returned with the sequence of band 25. This DNA species matched both MHV-68 ORF 58 and 59, which therefore suggested that they were transcribed as part of the same mRNA. Indeed subsequent sequence analysis (Milligan, 1998) has shown that ORF 58, 59 and 60 are likely to share a poly-A tail.

It should also be noted that the cells infected with MHV-68 in this study were firstly tissue culture adapted, i.e. immortalised, cells and also that they were derived from baby hamster kidney cells. How closely related the behaviour of tissue culture cells and cells in a living organism are, has been often conjectured. In particular, with respect to attempts made to highlight host genes that are up-regulated by the virus following infection, the merits of using a hamster tissue culture cell line to simulate small changes that are occurring in mice infected with MHV-68 could be argued. However, for the purposes of this study, namely setting up and optimising a differential display system to study MHV-68, using the well characterised BHK cell line in which the virus was known to replicate efficiently was perfectly satisfactory. For further study of MHV-68 infection *in vitro*, switching to a contact-inhibited murine cell line (such as NIH 3T3 cells) could be valuable to better simulate the *in vivo* environment.

In conclusion, a differential display system was set up and optimised for the study of MHV-68 infections *in vitro*. Data was produced that was independently confirmed using northern blots. The potential of the differential display system to identify qualitative differences has been amply demonstrated here. However, the quantitative aspect of the data was not so clear. Without a more exhaustive data set it is difficult to be certain but the differential display developed in this study was not able to consistently resolve relative differences in expression.

It could be that differential displays are best suited for looking at differences in the expression of a small subset of the transcript population between samples. Such small differences would be far easier to identify and increase in noise at the reamplification stage would be less of a hindrance than in the comprehensive analysis of transcriptional changes during infection that is being attempted in this study. Also, even in this small study, there was a level of redundancy appearing in

the data which would also be problematic for a comprehensive analysis but less so when looking for changes in a small number of genes.

As the differential display technique has been optimised, it could be used to examine different aspects of MHV-68 infection. As noted previously, there could be a smaller number of differentially expressed genes at very early stages of infection. Therefore, the differential display could be used for a focused study of these events. In particular, the host response at the start of infection could be very interesting to dissect transcriptionally.

However, for the purposes of this study, a different system for analysis of gene expression appears to be required. Relatively little is known of MHV-68 gene expression and function of those genes. A suitable technique requires that the transcription of a large number of genes can be examined simultaneously. Of the technologies that are accessible, arrays appear to fulfil the needs of this study better than the differential display system. Therefore, the differential display was put on hold and an array system was investigated.

4 Design and Development of an Array System for the Analysis of MHV-68 Gene Expression

4.1 DNA Array Design Strategy for MHV-68

The aim of the work presented in this chapter was to design and develop a global DNA array system for the analysis of MHV-68 gene expression. An array for the analysis of HHV-8 gene expression was used as a guide (Jenner, 2001). The chosen strategy consisted of arrays made by spotting cDNA probes onto nylon membranes. Hybridisations were performed with radiolabelled target reverse transcribed from total cellular RNA.

The following plan was devised to produce the MHV-68 arrays (see in Figure 4.1): Primers were designed (Primer3 software, Rozen, 2000) to PCR amplify the probe sequences from MHV-68 DNA. The resulting PCR products were cloned into plasmids to allow sequence verification, and also to provide a master template for subsequent amplifications. Once the probe sequences were confirmed, they were then reamplified from the relevant plasmid stocks, purified and quantified. The probes were spotted onto nylon membranes, and denatured prior to hybridisation.

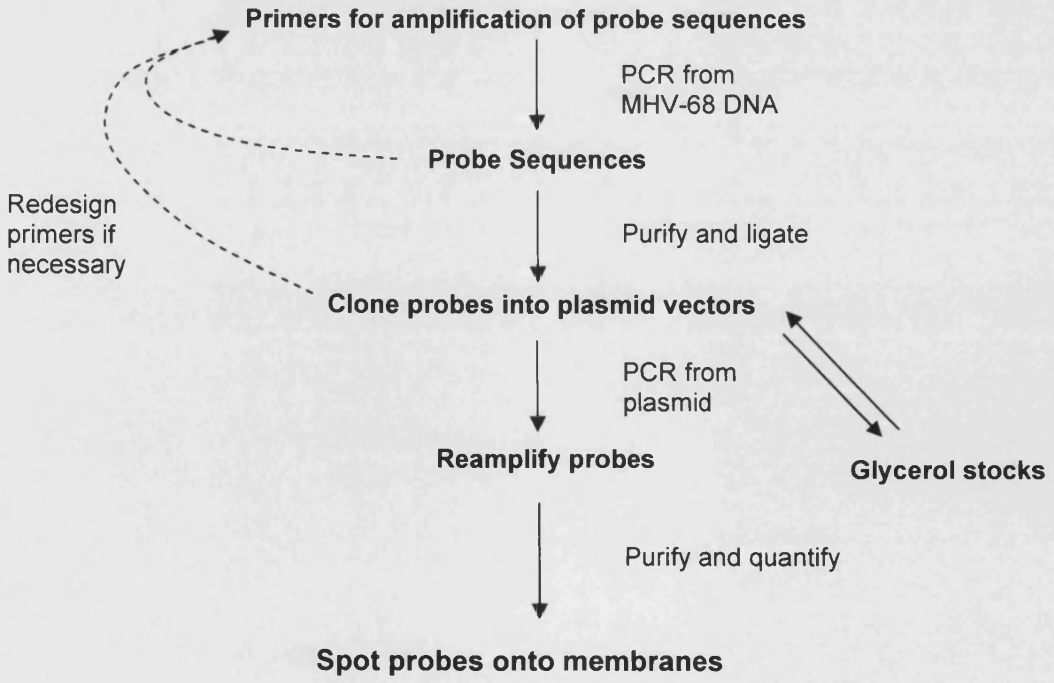


Figure 4.1. MHV-68 DNA array production process outline.

4.2 Design of the Probeset

All 80 predicted MHV-68 genes (Virgin, 1997) and any inter-genic regions larger than 100bp were represented in the probeset. Nine host-cell housekeeping genes were included to allow normalisation, and therefore inter-array comparisons. Three negative controls were also present to set non-specific hybridisation levels. Each probe sequence was approximately 300bp in size and corresponded to the 5' end of it's ORF.

4.2.1 Viral Probes

Primers for the amplification of viral probes were designed based on the published sequence data and genome analysis (Virgin, 1997). This analysis showed 80 MHV-68 genes, and so probes were designed to each of the 80 ORFs. For ease of identification, probes were arbitrarily named A1-A12, B1-B12, C1-C12, D1-D12, E1-E12, F1-F12 and G1-G8. A list of viral probes and the identities of each probe are shown in Table 4.1.

4.2.2 Inter-genic Probes

Initial analysis of the coding content of the MHV-68 genome had been performed using AceDB (a genome database system originally designed for the *C. elegans* genome project), and then individual ORFs were analysed by BLAST, allowing comparison with other herpesviruses (Virgin, 1997). Therefore, all MHV-68 gene designations were based purely on bioinformatics and not experimental evidence.

This left the possibility that some MHV-68 genes could have been missed. Therefore, probes were also designed to inter-gene regions, whenever they were larger than 100bp in size. Furthermore, to cover the possibility of coding strands being present on either strand of the MHV-68 genome, these inter-genic probes were designed in both orientations.

There were 6 such inter-genic regions, resulting in 12 probes designated G9-G12 and H1-H8.

4.2.3 Housekeeping and Negative Control probes

Nine housekeeping genes were selected due to their maintained expression in a number of physiological states and their universal adoption for this role in a number of similar systems (Adams, 1995; Clontech, 1998). Therefore, probes were designed for glyceraldehyde-3-phosphate dehydrogenase (GAPDH), myosin 1, murine ornithine decarboxylase (MOD), β-actin, calcium binding protein 45 (Cab45), ribosomal protein S29, ubiquitin, phospholipase A2 and hypoxanthinie

phosphoribosyl transferase (HPRT). The probes of these housekeeping genes were designated H9-12 and J1-5.

Three negative controls were also included to assess non-specific binding of target to probes: pBluescript II (SK+) plasmid (Stratagene), TMV 180kDa protein and water. The TMV probe was designated J6. The pBluescript plasmid was spotted directly onto membranes.

4.3 Production of Probe Sequences to be Spotted onto Arrays

4.3.1 Primer Design

The primers for amplification of probe sequences were designed using Primer3 software (Rozen, 2000). Every primer was designed to have the same properties: Primer size 18-22bp, optimal 20. Primer melting temperature 58-62°C, optimal 60°C. Primer GC% 40-60%, optimal 50%. 3' GC clamp 2bp. Optimal product size 300bp. Template sequence was the viral genome. For the housekeeping genes, the mRNA sequences for the respective genes were used. All other properties were set to the default settings in the software. Primers for the TMV probe were kindly provided by Dr Paul Kellam (Wohl Virion Centre, UCL).

To minimise any potential non-specific hybridisations between probes and targets, the probe sequences were entered into a database and the BLAST algorithm used to check for potential cross-hybridisations. The same process was repeated with the primers used to amplify the probes, as they would be used together in the array primer mix, to prime the target synthesis and labelling reaction. None of the probe or primer sequences showed significant homology to each other.

Occasionally the primer design conditions could not be met due to the presence of overlapping genes, or genes shorter than 300bp. As each probe had to be gene-specific and therefore unique, this required the design conditions to be changed in these cases. For example the MHV-68 ORF M10a, M10b and M10c are overlapping. Any part of a ORF that overlapped another ORF was removed from the template sequence used for primer design. This meant that the M10c probe was only 196bp in size as there was only around 200bp of unique sequence for M10c. ORF 53, M12 and M13 also had probes smaller than 300bp for the same reasons. Unfortunately, ORF M10a or M10b had no regions of unique sequence and therefore no probes were designed for these ORFs. A table of primer sequences can be found in Appendix I. Table 4.1 lists the probes and their product sizes.

Table 4.1. DNA Array probes and their sizes.

Designation	Gene	Probe Size
A1	M1	320
A2	M2	301
A3	M3	305
A4	M4	301
A5	ORF4	318
A6	ORF6	285
A7	ORF7	291
A8	ORF8	316
A9	ORF9	304
A10	ORF10	299
A11	ORF11	300
A12	K3	281
C1	ORF27	300
C2	ORF29b	303
C3	ORF30	222
C4	ORF31	285
C5	ORF32	299
C6	ORF33	323
C7	ORF29a	288
C8	ORF34	314
C9	ORF35	309
C10	ORF36	301
C11	ORF37	304
C12	ORF38	210
E1	ORF52	285
E2	ORF53	170
E3	ORF54	320
E4	ORF55	318
E5	ORF56	283
E6	M8	301
E7	ORF57	292
E8	ORF58	308
E9	ORF59	308
E10	ORF60	311
E11	ORF61	289
E12	ORF62	320
G1	ORF73	298
G2	ORF74	307
G3	ORF75c	318
G4	ORF75b	301
G5	ORF75a	304
G6	M12	147
G7	M13	99
G8	M14	275
G9	M1-M2	292
G10	M2-M1	313
G11	M2-M3	301
G12	M3-M2	309
J1	Cab45[3]	314
J2	R S29[4]	219
J3	Ubiquitin	296
J4	Phospholipase A2	291
J5	HPRT[5]	304
J6	TMV[6]	

Designation	Gene	Probe Size
B1	M5	320
B2	M6	303
B3	ORF17	311
B4	ORF18	308
B5	ORF19	298
B6	ORF20	320
B7	ORF21	310
B8	ORF22	316
B9	ORF23	317
B10	ORF24	292
B11	ORF25	300
B12	ORF26	306
D1	ORF39	320
D2	ORF40	304
D3	ORF42	289
D4	ORF43	301
D5	ORF44	288
D6	ORF45	287
D7	ORF46	298
D8	ORF47	294
D9	ORF48	290
D10	ORF49	301
D11	ORF50	307
D12	M7	318
F1	ORF63	293
F2	ORF64	302
F3	M9	283
F4	ORF66	294
F5	ORF67	320
F6	ORF68	293
F7	ORF69	298
F8	M10a	None
F9	M10b	None
F10	M10c	196
F11	ORF72	310
F12	M11	309
H1	M3-M4	311
H2	M4-M3	298
H3	K3-M5	293
H4	M5-K3	289
H5	ORF27-29b	154
H6	ORF29-27	154
H7	M10c-ORF72	313
H8	ORF72-M10c	313
H9	GAPDH[1]	293
H10	Myosin 1	301
H11	MOD[2]	299
H12	β-actin	309

[1]Glyceraldehyde-3 phosphate dehydrogenase
[2]Murine ornithine decarboxylase
[3]Calcium binding protein 45
[4]Ribosomal protein S29
[5]Hypoxanthinie phosphoribosyl transferase
[6]Tobacco mosaic virus 180kDa protein

4.3.2 MHV-68 DNA Isolation

To provide template for PCR amplification of probe cDNAs, MHV-68 DNA was isolated from 200µl of purified MHV-68 stock and the integrity of the resultant DNA was tested by PCR for a number of the larger viral genes (data not shown). This showed that the isolated viral DNA was of sufficient quality to act as template for the amplification of 300bp cDNA probes.

4.3.3 PCR Amplification of Viral Probe Sequences

High yields of cDNA were preferable as large stocks of cDNA were required for spotting onto arrays. Reaction conditions for the amplification of probe cDNA were optimised. Varying the amount of template DNA was found to effect yield (Figure 4.2): Amplifications with higher template concentrations of DNA produced smeared bands, however smearing below 300bp was removed when lower DNA template concentrations were used. Also a high molecular weight band, possibly template DNA, was observed at the highest template concentrations (lanes 1 and 2, Figure 4.2). Diluting out MHV-68 DNA template resulted in the loss of this high molecular weight band. Therefore, the MHV-68 DNA working stock was set at 1/1000 dilution of original stocks.

Initially, PCR reactions were performed using the following conditions: 1x PCR buffer (Invitrogen), 1.5mM $MgCl_2$, 1mM dNTPs (Promega), 2µM pGEM-T Easy vector-specific primers, 50U/ml Taq polymerase (Invitrogen) and 1µl working stock MHV-68 DNA as template in a final volume of 25µl. This reaction mix was incubated for 5min at 95°C, followed by 30 cycles of 95°C (30s), 58°C (30s) and 72°C (30s), before a final elongation of 5min at 72°C. The reaction was then held at 4°C, before products were analysed by agarose gel electrophoresis.

Ninety viral probe sequences (ORF and inter-gene) were amplified under these conditions, as listed in Table 4.1. The reaction products were electrophoresed on an agarose gel as shown in Figure 4.3. Seventy-two probe sequences were successfully amplified in this way, although some of these successful reactions showed low yields. The remaining 18 reactions produced multiple products, products of the wrong size or no product at all.

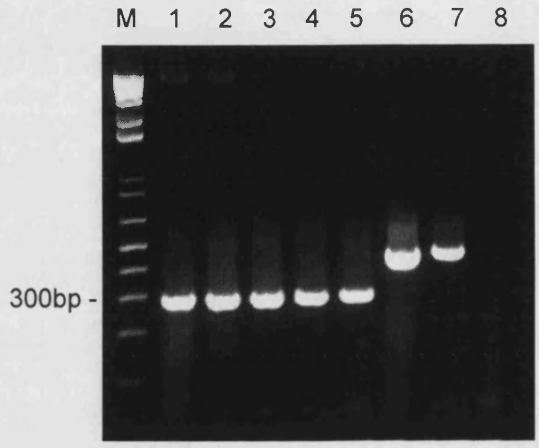


Figure 4.2. Agarose gel analysis of PCR amplification of viral genes using MHV-68 DNA template.

Lane M shows the molecular weight marker. Lanes 1-5 show the results of amplification with ORF 2 primers, with reducing template concentrations: 1/100, 1/200, 1/500, 1/1,000, 1/10,000 dilutions of the stock, respectively. Lanes 6 is a positive control: ORF25 was amplified with tried and tested primers from MHV-68 DNA. Lane 7 is a control for the MHV-68 DNA template. Again ORF 25 was amplified as in lane 6, but the template used was cDNA that had been reverse transcribed from RNA isolated from MHV-68 infected NIH 3T3 cells. Oligo-dT was used to prime the reverse transcription. Lanes 8 is a negative control in which no DNA template was included in the reaction mix.

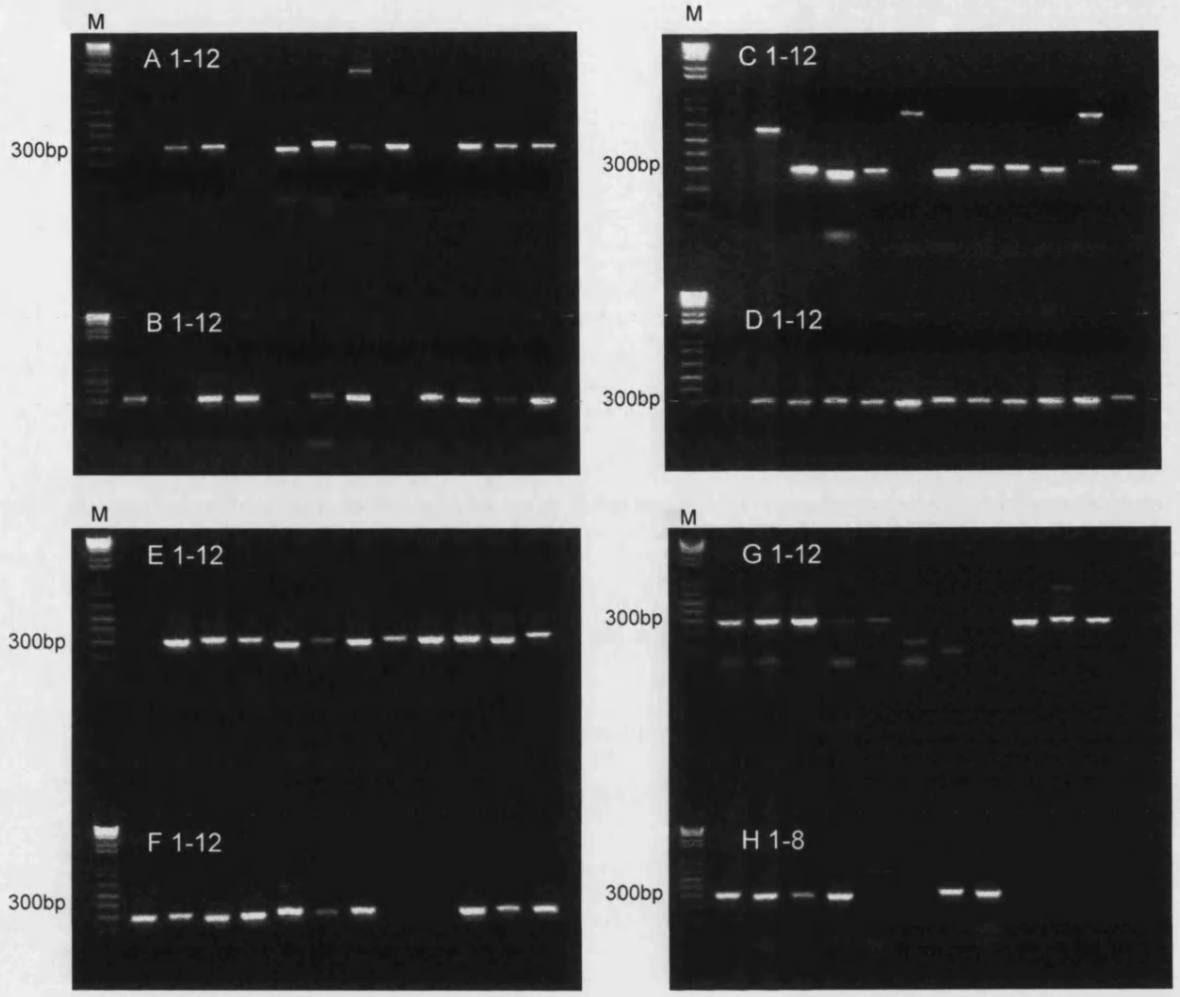


Figure 4.3. PCR amplification of viral DNA array probe sequences.
Probe sequences were amplified from MHV-68 DNA, using an annealing temperature of 58°C. Individual probes are identified in Table 4.1.

Amplification of probes A7, C2, C6, C11 and H5 resulted in multiple reaction products, or a product of the wrong size. This suggested that the amplification had been non-specific. Therefore, amplification of probe sequences was repeated with an increased annealing temperature of the 60°C, which would increase the specificity of primer binding to template. This resulted in the successful amplification of 3 sequences: probes A7, C2 and C11.

Reactions producing no product were repeated with reduced annealing temperatures of 56°C and 54°C. Also the effects of increasing and reducing template concentration were evaluated. However, none of these modifications resulted in successful amplification of product.

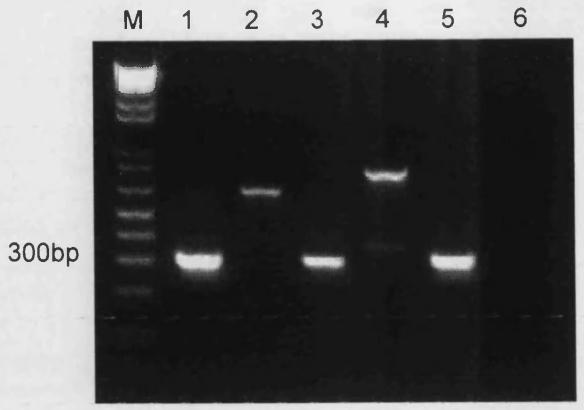


Figure 4.4. PCR amplification of viral probe sequences with increased annealing temperature.

PCR amplification of probe sequences was repeated with an annealing temperature of 60°C. Lanes 1-8 represent probes A7, B6, C2, C6, C11, & H5, respectively. M denotes the marker lane.

For the remaining 15 probe sequences (A1, A4, A9, B2, B5, B8, C1, C6, E1, F8, F9, G8, G12, H5, H6) new primers were designed. Repeating the amplification with the new primers was successful for all remaining probe sequences, except G8. The agarose gel electrophoresis is shown in Figure 4.5. It should be noted that H5 (lane 14) and H6 (lane 15) are the expected sizes, as there was only approximately 150bp of unique sequence for these probes.

For G8, the PCR amplification was repeated varying the annealing temperature, $MgCl_2$ concentration and template concentration. However, none of these changes allowed probe G8 to be amplified successfully. Therefore, a third set of primers were designed and used for PCR amplification reactions, but again without success. Examining the sequence data of the G8 gene (MHV-68 ORF M14) showed that the sequence contained a high proportion of GC bases. Therefore, the PCR reaction was repeated with the inclusion of DMSO, as certain sulphoxides are well known to aid in the amplification of GC-rich templates (Bookstein, 1990; Pomp, 1991). In addition, an alternate, high efficiency, heat activated Taq polymerase was also tested (without DMSO; Advantage Taq, Clontech). The agarose gel electrophoresis of reaction products is shown in Figure 4.6. Inclusion of 10% DMSO in lane 1 proved successful and although 2 bands resulted, 1 was the correct size and therefore excised and purified. Using Advantage Taq for PCR amplification, shown in lane 3, proved unsuccessful.

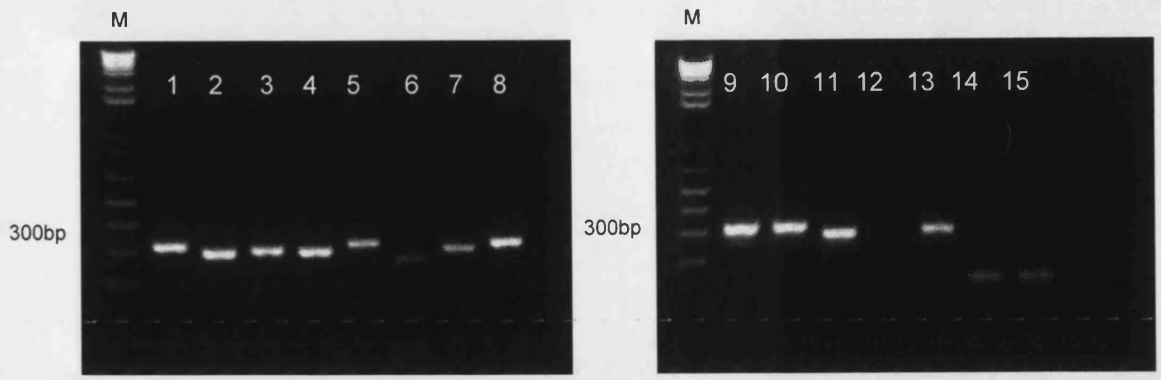


Figure 4.5. PCR amplification of further viral probe sequences.
PCR amplification was repeated using new primers. Lanes 1-15 show A1, A4, A9, B2, B5, B8, C1, C6, E1, F8, F9, G8, G12, H5 & H6, respectively.

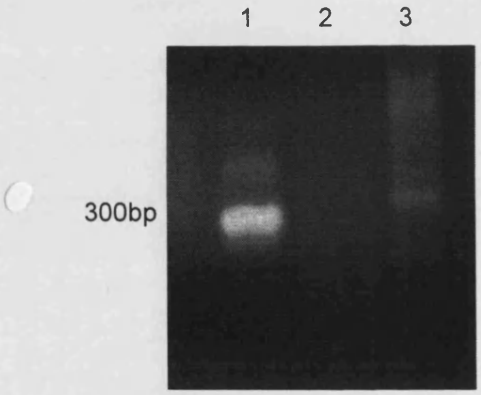


Figure 4.6. Amplification of probe G8 sequence.
PCR amplification of G8 using with 10% and 20% DMSO in lanes 1 & 2, respectively. Lane 3 shows a PCR reaction using Advantage Taq polymerase (Clontech) and no DMSO.

4.3.4 PCR Amplification of Control Probe Sequences

RNA was isolated from NIH 3T3 cells and reverse transcribed with anchored oligo-dT primers. The resulting cDNA was then used as template for PCR amplification of the control probes to feature on the array as shown in Figure 4.7. All housekeeping genes, as listed in Table 4.1, were successfully amplified.

A pGEM-T Easy plasmid containing the TMV probe was kindly donated by Dr P. Kellam (Wohl Virion Centre, UCL). Bluescript plasmid was obtained from Stratagene for spotting directly onto array membranes.

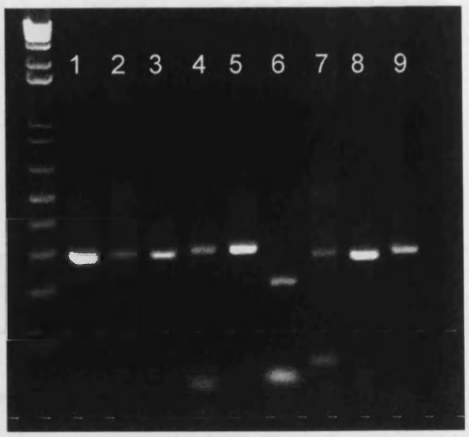


Figure 4.7. PCR amplification of DNA array control probe sequences.
Control probe cDNA were amplified by PCR from 3T3 cell cDNA using the appropriate primers (Table 4.1). Lanes 1-9 represent probes H9-12 and J1-5, respectively.

4.3.5 Cloning of Probe Sequences

All probe sequences were cloned into the pGEM-T Easy system for sequence verification. This system allowed direct ligation of purified PCR products into the pGEM-T cloning vector. After transformation, the *E. coli* JM109 culture was plated out for colour screening. Successful ligation of an insert into cloning vector resulted in the disruption of a β-galactosidase coding region. By plating the transformed bacteria on agar plates containing X-gal and IPTG, presence of an insert was indicated by white as opposed to blue colonies. Positive colonies were picked, cultured and glycerol stocks frozen down.

Glycerol stocks were checked for presence of the correct probe sequence insert by PCR. Both the presence of the correct insert and the presence of a single insert was checked using insert-specific (see Appendix I for sequences) and plasmid-specific primers (see 2.3.19 for sequences). The plasmid-specific primers were designed to amplify any sequence ligated into the multiple cloning site of the pGEM-T vector. Any mis-clones were discarded and the relevant probe sequence re-cloned. All probe sequences were successfully cloned.

As a precaution, cultures were prepared from the glycerol stocks, and plasmid isolated. Each plasmid was verified to contain the correct insert. This ensured that the glycerol stocks were viable and confirmed that the probe sequences had been successfully cloned.

4.3.6 Verification of Cloned Probe Sequences

Bacterial cultures were prepared from glycerol stocks and plasmid isolated. These plasmids were sequenced in house to verify the insert. The sequence data for each plasmid was used to search nucleotide databases with BLAST. Any plasmids that contained the wrong insert were discarded.

All plasmids were sequence verified to contain the appropriate insert, except probe G5 (MHV-68 ORF 75a) and G8 (MHV-68 ORF M14). PCR amplification of these plasmids using probe-specific primers resulted in a band of the expected size; however the plasmid could not be sequenced. Due to the GC-rich nature of the inserts DMSO, was added to the sequencing reaction but this also failed. Even commercial sequencing by MWG Biotech was unsuccessful. These 2 probes were incorporated into the array but with the caveat that these probes were not sequence verified.

4.3.7 Amplification of Probe Sequence cDNA from Plasmid

Large amounts of probe sequence cDNA were required for spotting onto arrays. A standard high efficiency PCR amplification was set up to allow efficient amplification of cDNA from the plasmids for each probe. Plasmid-specific primers were used in an optimised protocol (see Chapter 2.4.20) that used high concentrations of Taq polymerase, primers and deoxynucleotides: 1x PCR buffer, 1.5mM $MgCl_2$, 2mM dNTPs, 4μM pGEM-T Easy vector-specific primers, 50U/ml Taq polymerase and 30μg/μl plasmid template in a final volume of 100μl. This reaction was incubated for 5min at 95°C, followed by 40 cycles of 95°C (30s), 55°C (30s) and 72°C (30s), before a final elongation of 5min at 72°C. The reaction was then held at 4°C.

4.3.8 Quantification of Probe cDNA

Three methods for the quantification of the probe cDNA were used: Fluorescence of DNA spotted onto ethidium bromide containing agarose gels, weight-standardised DNA hyperladders and spectrophotometry.

To test ethidium bromide-based quantification, known quantities of salmon sperm DNA were spotting onto an ethidium bromide containing agarose gel and the fluorescence under UV light was measured. Plotting these measurements returned the standard curve shown in Figure 4.8a, with most points lying within the 95% confidence limits, and a correlation coefficient of 0.93. However, this method was found to give inconsistent results when a second set of standards were spotted randomly and the calculated values compared to actual values, as shown in Figure 4.8b. Therefore, ethidium bromide quantification was unsuitable for quantification of

amplified probe sequences. DNA hyperladders also produced inconsistent results and were therefore also unsuitable.

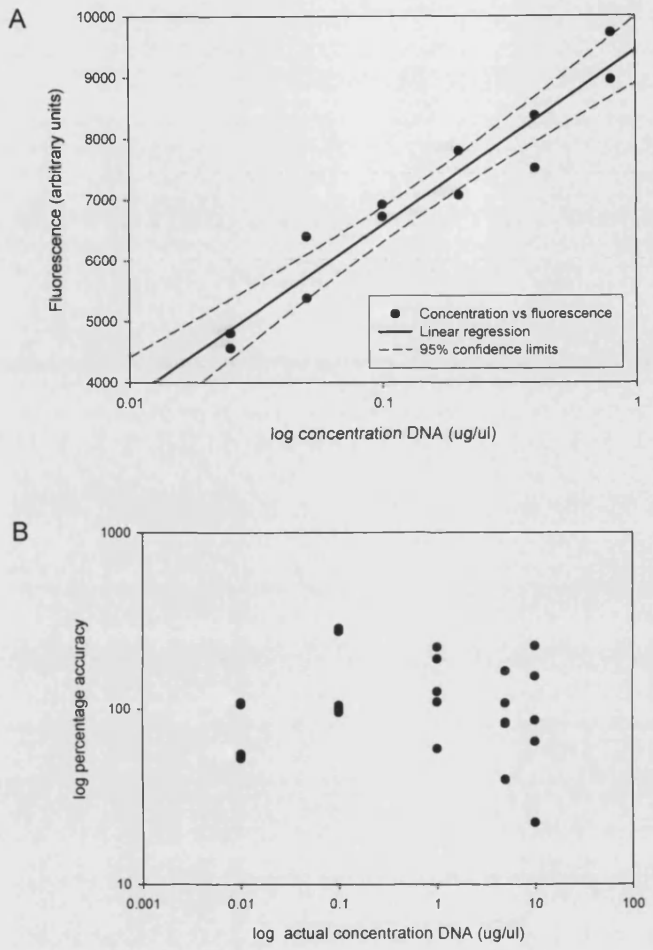


Figure 4.8. Ethidium bromide quantification of DNA.

A. Known quantities of salmon sperm DNA were spotted onto an agarose gel containing ethidium bromide. DNA quantity was plotted against fluorescence and linear regression used to produce a standard curve. B. Consistency of ethidium bromide quantification. Known quantities of DNA were spotted randomly onto an agarose gel containing ethidium bromide. Each spot was quantified using the standard curve derived in A. The calculated values were plotted as a percentage of their actual values. Each quantity of DNA was spotted 5 times.

Although spectrophotometry had the disadvantage of requiring large amounts of DNA relative to the other methods, it was the most accurate and reliable one. Therefore, this method was used to quantify probe sequence DNA. Additional PCR amplifications of probe cDNAs were performed to provide enough DNA for quantification, as DNA that had been quantified was no longer suitable for spotting

onto arrays. Amplifications of probe sequences from plasmid, using the optimised protocol, were found to produce approximately 5µg of cDNA per reaction. Once quantified, all probe cDNAs were diluted to 0.2µg/µl and stored at -20°C.

4.4 Array Manufacture and Optimisation of the System

4.4.1 Initial Array Protocol

In order to test array designs, an initial array hybridisation protocol was necessary. An overview is shown in Figure 4.9. The following protocol was used: NIH 3T3 cell monolayers were infected with MHV-68 at moi 10. Control cells were mock-infected. At the appropriate time pi, the monolayers were lysed with TRIzol and total RNA isolated. The RNA was treated with DNAse I to remove any contaminating DNA. Each RNA sample was quantified and run out on a denaturing gel to assess quality.

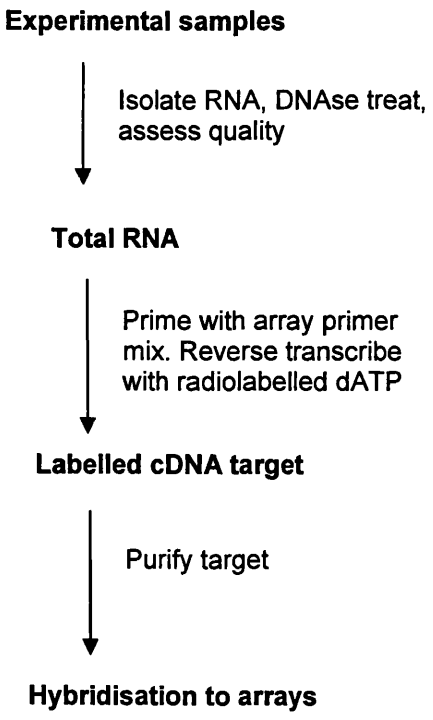


Figure 4.9. Overview of DNA array protocol.

Total RNA was used as template for the labelling reverse transcription reaction. Labelled target was synthesised by reverse transcription, with the inclusion of ^{33}P-dATP. The reaction was primed with a primer mix consisting of all the 3' primers that had been used to amplify probe sequences. The final concentration of each

primer in the reaction was 0.1μM. The reaction was terminated by the addition of EDTA.

The labelled pool of target cDNA was purified to remove unbound ^{33}P-dATP and hybridised to arrays. Arrays were prehybridised for 1h at 60°C in Expresshyb and sheared salmon sperm DNA. Hybridisation was performed in fresh Expresshyb and sheared salmon sperm DNA at 60°C for 16h. Arrays were then washed 3 times at low stringency, followed by 3 washes at high stringency.

4.4.2 Testing Membranes

Before manufacturing whole arrays, the membrane to be used was tested. The 2 membranes tested were uncharged nylon and positively charged nylon. A number of probes were spotted onto the uncharged nylon (Biorad) and on to the positively charged nylon membrane (Amersham).They included MHV-68 ORF M3, 53, 67 and housekeeping gene probes GAPDH, myosin and phospholipase. Both arrays were used in a test hybridisation, as shown in Figure 4.10. It was clear that the positively charged membrane in A performed much better than the uncharged membrane in B. Therefore, the positively charged nylon membrane was chosen for making all subsequent arrays.

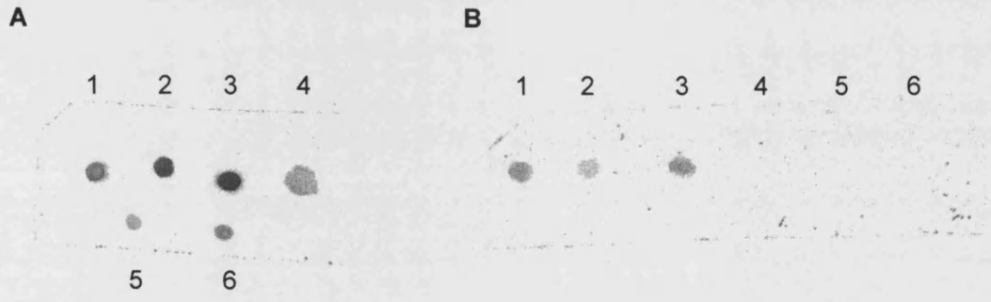


Figure 4.10. Comparison of array membranes.
A. Negatively charged nylon membrane (Amersham). B. Uncharged nylon membrane (Biorad). MHV-68 probes for ORF M3 (1), 53 (2), 67 (3) and housekeeping gene probes D-glyceraldehyde -3-phosphate dehydrogenase (4; GAPDH), myosin (5) and phospholipase (6) were spotted using a gilson pipette. Labelled target cDNA was hybridised to both membranes and then visualised on a phosphoimager.

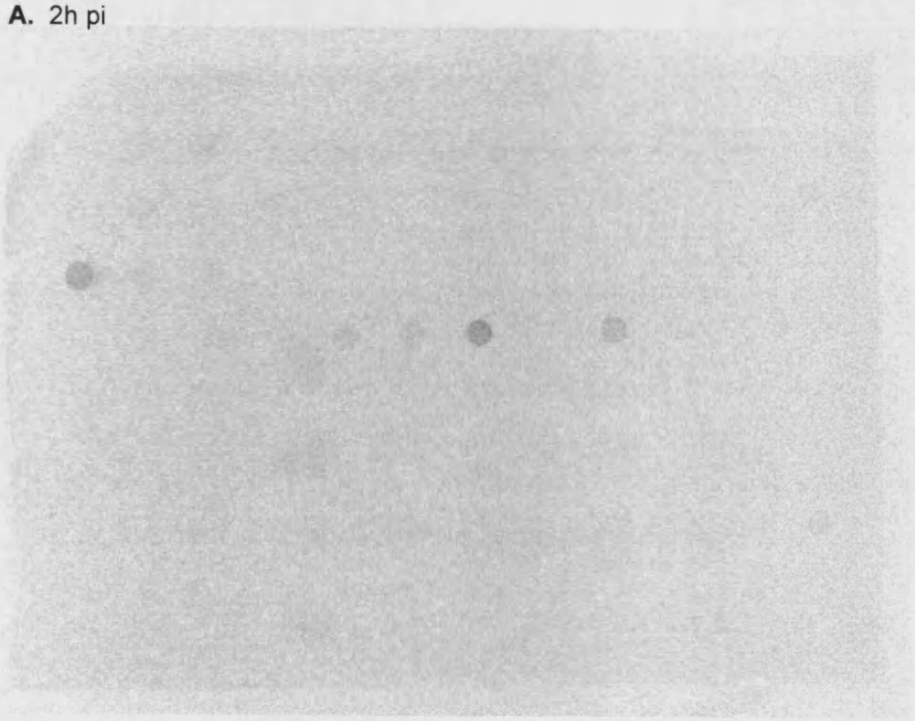
4.4.3 Initial Hybridisation

Arrays were produced by vacuum-spotting 400ng of each probe DNA onto membranes. The layout of the 96 spot vacuum-spotted array is shown in Appendix II. The top left corner of each array was cut to aid orientation. An initial test hybridisation to a 96 spot vacuum-spotted array using the initial array protocol is shown in Figure 4.11. Overall the array showed weak hybridisation between probes and targets. This suggested that hybridisation efficiency was low, due possibly to insufficient labelled cDNA target, insufficient probe, or suboptimal hybridisation conditions. Furthermore, the control probes for housekeeping genes (present in the bottom row of spots on each array) did not appear to bind target.

4.4.4 Optimising Synthesis and Labelling of Target cDNA

Increasing the quantity of target cDNA was attempted by optimising reaction conditions. A number of target synthesis reactions were set up to assess the effects of varying the concentrations of constituents and also varying the reaction conditions themselves. (This optimisation was performed without radiolabelled dATP to minimise the use of radioisotopes.) To increase the efficiency of priming, RNA was incubated with the array primer mix for 5min at 70°C and then placed on ice, prior to addition of the reaction mix.

A. 2h pi



B. 24h pi

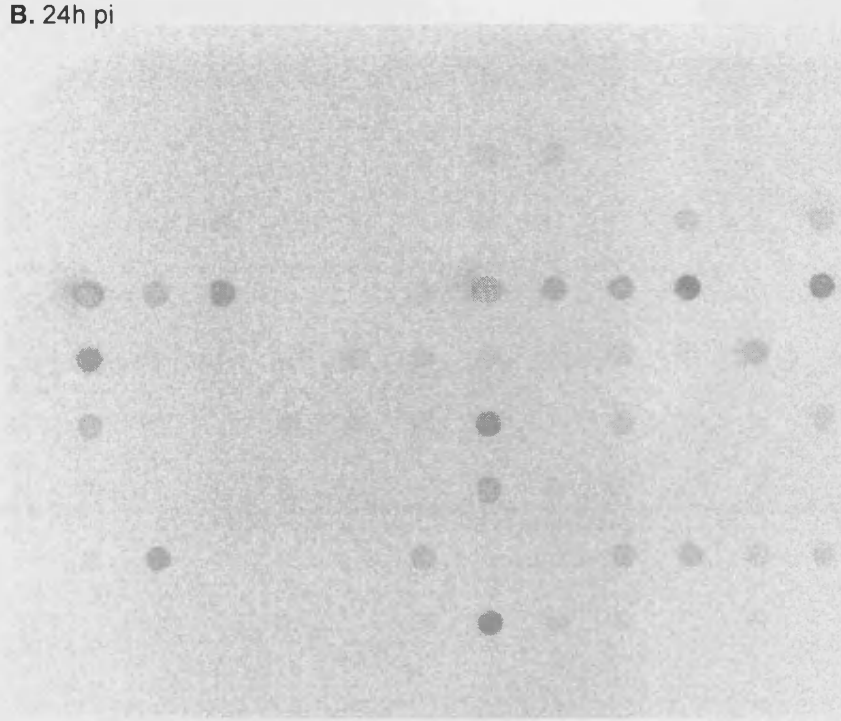


Figure 4.11 (this and preceding page). Testing vacuum-blotted DNA arrays. Hybridisations with labelled cDNA target derived from NIH 3T3 cells infected for 2h and 24h are shown here. A key to the layout of the array is shown in Appendix II.

Replacing MMLV reverse transcriptase with Superscript II reverse transcriptase was found to improve the yield of cDNA, as a longer incubation at a higher temperature was possible. This was due to Superscript II lacking the RNAse H activity usually associated with reverse transcriptases.

Increasing primer concentration was also found to improve the yield of cDNA, which suggested that insufficient primer had been included in the initial target synthesis reactions.

Therefore, optimal yields for target cDNA synthesis were found to be achieved with the following conditions:

1. Firstly, template RNA was annealed to primers for 5min at 70°C, before being placed on ice, to increase the efficiency of priming.
2. The reaction mix was then added and consisted of 1x first strand buffer, 1mM dA/T/G/CTP and 5U/μl Superscript II reverse transcriptase, in a final volume of 20μl. The reaction mix was then incubated at 50°C for 90min. For actual labelling reactions the dATP was replaced with 2.5μCi/μl [α-^{33}P] dATP.

4.4.5 Optimising Purification of Labelled Target

The following methods of purifying labelled cDNA were assessed: NICK columns (Amersham), Chromaspin columns (Clontech) and Nucleospin columns (Clontech). The Chromaspin and Nucleospin columns achieved the best results as measured by ability to remove 20bp oligonucleotides whilst retaining DNA fragments 100bp and greater in size. The Nucleospin columns were the more practical choice, as they could be used in a standard benchtop centrifuge. They were therefore selected as the method of choice.

4.4.6 Optimising Hybridisation Conditions

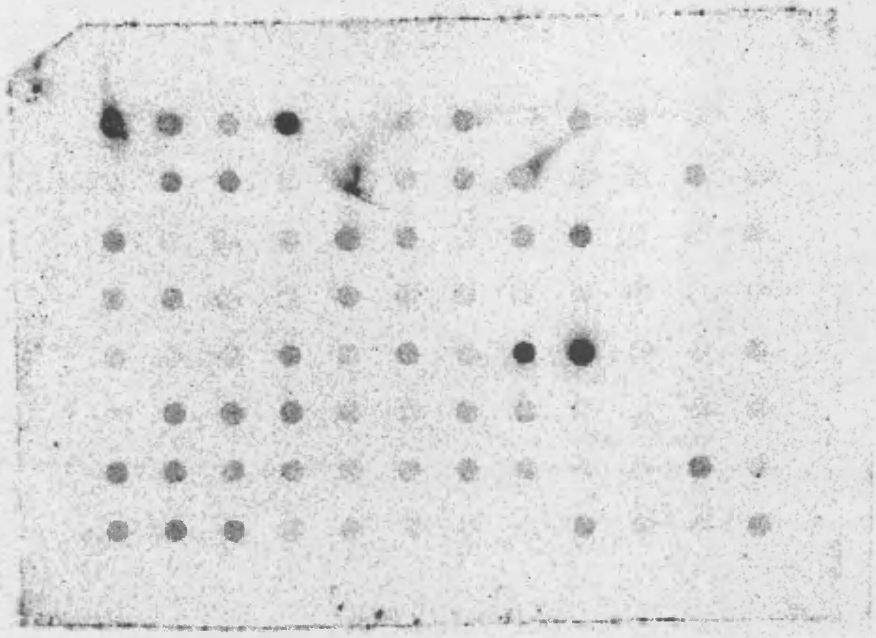
A second hybridisation was performed, using the protocol as optimised so far. RNA was isolated from cells infected for 2h and 20h, as well as from a mock-infected control. The results of the hybridisation are shown in Figure 4.12. Although signal strengths had increased compared to the initial hybridisation, there was little difference between the hybridisation patterns of infected and uninfected hybridisations. Labelled target had hybridised to viral probes even in the hybridisation using target derived from mock-infected cell RNA. (The housekeeping gene probes on these arrays were along the bottom row, and the rest of the array consisted of viral probes, as shown in Appendix II). This suggested that the results of these hybridisations were due largely to non-specific binding, as opposed to true signals.

The arrays were then subjected to additional washes at high stringency, to establish whether inefficient washing has been responsible for the non-specific signals seen in Figure 4.12. However, these further washes had no effect. The RNA samples were checked for presence of viral mRNAs by RT-PCR, which confirmed that the mock-infected RNA sample did not contain viral RNAs.

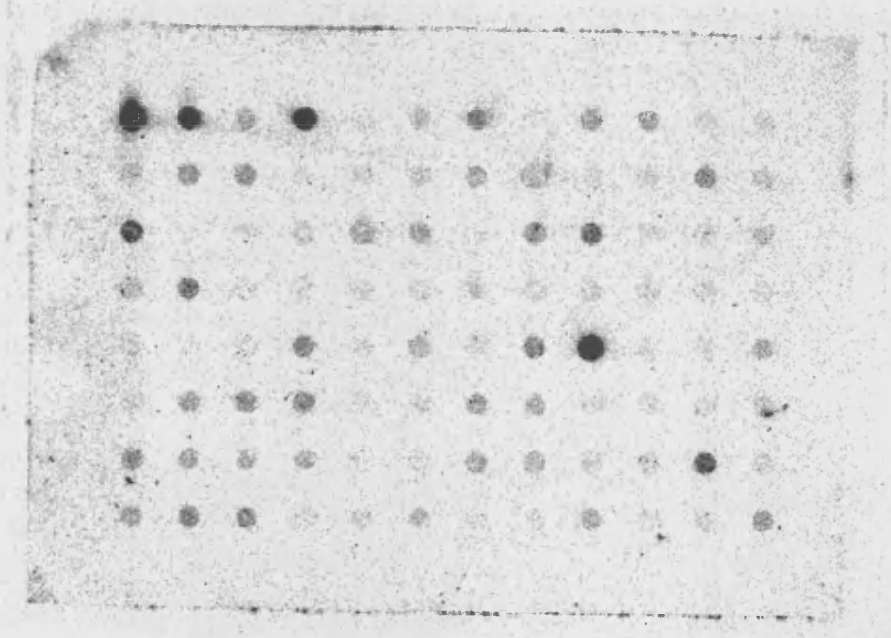
4.4.7 Use of Miniarrays for Further Optimisation of Hybridisation Conditions

To further optimise the system, vacuum-spotted miniarrays were made. Each consisted of a small membrane with 6 viral and 2 housekeeping probes (MHV-68 ORF M6, 22, 43, 52, 57, 61 and β-actin, GADPH). These were used as multiple array hybridisations are easier to setup with smaller arrays, and the analysis of 8 probe signals was far simpler than analysing full arrays.

A. Mock infected



B. 2h pi



C. 20h pi

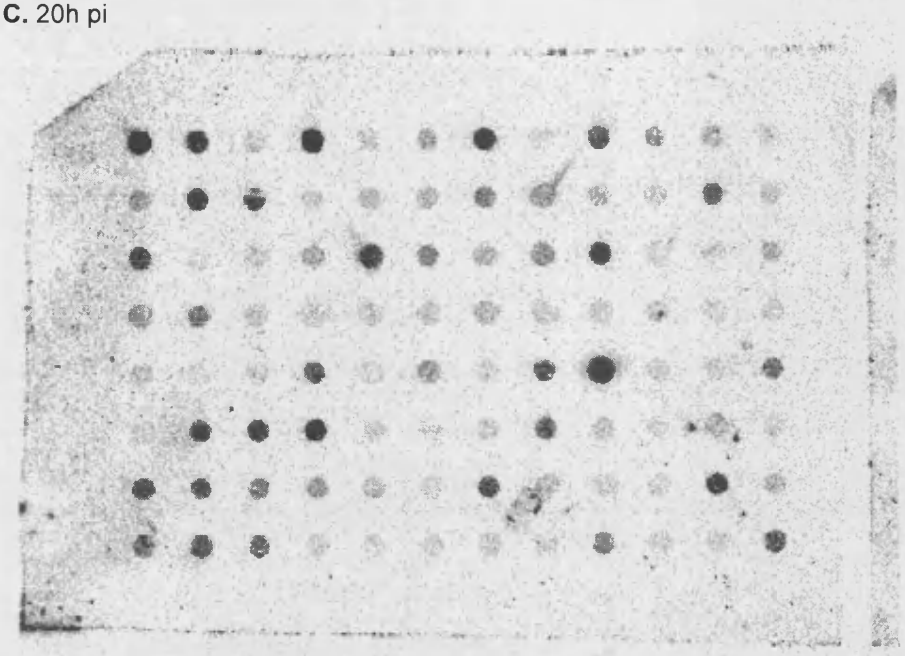


Figure 4.12 (this and preceding page). Hybridisation to vacuum-spotted MHV-68 DNA arrays.
Hybridisations were performed with labelled cDNA derived from mock-infected cells (A.), cell infected for 2h (B.) and for 20h (C.). A key the layout of the array is shown in Appendix II.

Hybridisations with varying conditions (temperature, time and hybridisation chamber rotation speed) were performed with these miniarrays to observe their effects on non-specific binding between targets and probes. However, no significant reduction of non-specific binding was observed by changing these hybridisation conditions.

The amount of RNA used as template was doubled to 4µg to see if there was a lack of specific binding due to lack of target cDNAs, but this also had little effect. The results of this hybridisation are shown in Figure 4.13.

4.4.7.a Optimisation of Probe Concentration

A second set of miniarrays were made to test the effects of changing the amount of probe in each spot on the array. Four probes were vacuum-blotted onto membranes in 3 quantities each: 800ng, 400ng and 200ng. Hybridisations using these miniarrays showed that spots with higher concentrations of probe gave stronger signals. Therefore, the amount of probe in each spot was a significant factor in the resulting signal strengths. Nonetheless, background levels were still high at all probe concentrations, as seen with hybridisations of uninfected sample-derived cDNA.

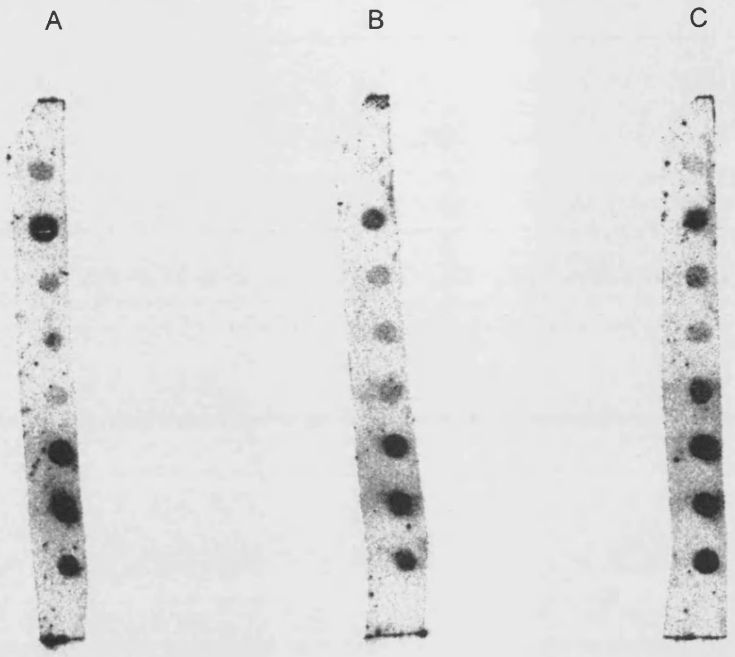


Figure 4.13. Hybridisation to DNA miniarrays.

Each miniarray was spotted with 6 viral and 2 murine gene probes: GADPH, β-actin and MHV-68 ORF 61, 57, 52, 43, 22 and M6 from top to bottom, respectively. Miniarrays had The top left corner of each miniarray was cut to aid orientation. A. Hybridisation to labelled target derived from uninfected cells. 2μg RNA template was used. B. Hybridisation to labelled target derived from cells infected for 12h. 2μg RNA template was used. C. Hybridisation to labelled target derived from cells infected for 12h. 4μg RNA template was used.

4.4.8 Summary of Optimising Hybridisation Conditions Using Vacuum-Spotted Arrays

Optimisation of hybridisation conditions had proved to be more complex than optimisation of target synthesis. In particular the results were often found to be inconsistent when repeating these experiments. This suggested that the arrays were not reproducible at this stage. Varying individual hybridisation conditions such as temperature, duration, speed of rotation and stringency of washing all had little effect in reducing non-specific binding of target to probe. The only factor identified to reduce background hybridisation levels consistently, was the addition of murine C_0T-1 DNA fragments to the hybridisation solution. However, this alone was not sufficient to allow considerable differences to be observed between infected and

uninfected samples. The amount of probe present in each spot was also found to influence binding of target, although its effects on non-specific binding were unclear.

Although many facets of the array had been optimised or attempts made to optimise them, the actual method of spotting the arrays had not been examined thus far. However, changing the concentration of probe in each spot was known to affect binding. Therefore, the method of spotting arrays, and the nature of the spots themselves was modified.

4.4.9 Pin-Spotted Arrays

Pin-spotting uses a very different mechanism to produce arrays, compared to vacuum-spotting. A fixed volume of probe solution is held as a droplet at the end of a pin. The droplet is then placed onto a dry nylon membrane by bringing the two into contact. With vacuum-spotting, the probe solution is drawn through a wet membrane under vacuum. The nature of the nylon membrane results in probe cDNA being left on the membrane, while smaller molecules pass through. Also the 2 methods differed in the denaturation of probe cDNA. The probes for vacuum-spotted arrays were denatured prior to spotting, and then cross-linked under UV light. Pin-spotted arrays were stored prior to denaturation of probes, which was instead performed just before use.

Pin-spotted arrays use a gentler mechanical force than vacuum-spotted ones to produce arrays. Also the spots produced by pin-spotting are approximately 1mm in diameter, around one third of that for vacuum-spotting. Therefore, very different mechanics of binding potentially exist for the two types of spot.

4.4.10 Biomek Pin-Spotted Arrays

As the amount of probe to be spotted onto membranes had yet to be optimised, arrays were produced with each probe spotted at 4 concentrations: 100ng/µl, 50ng/µl, 25ng/µl and 12.5ng/µl. As it was available, a Beckman Biomek 2000 robot was used to produce these. As this robot was capable of spotting 1,152 features onto a single membrane, each probe was spotted 3 times so that replicate probes featured on the array. The layout of this array is as shown in Appendix II. The bottom right corner of each array was cut to aid orientation.

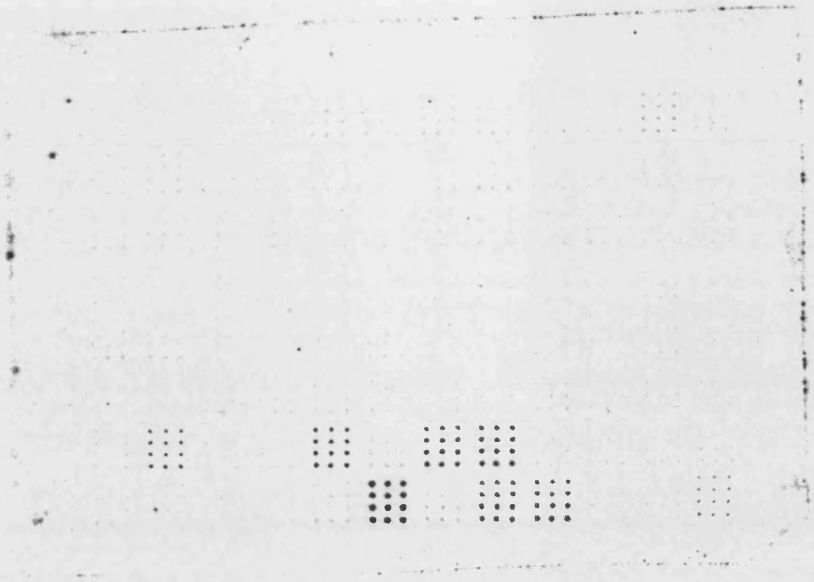
A hybridisation was performed using target derived from NIH 3T3 cells infected with MHV-68 for 12h, as well as from a mock-infected control, to assess the effects of using pin-spotted arrays. The optimised protocols were followed for target cDNA

synthesis and purification. The standard hybridisation and wash protocol was followed as used for the initial hybridisation.

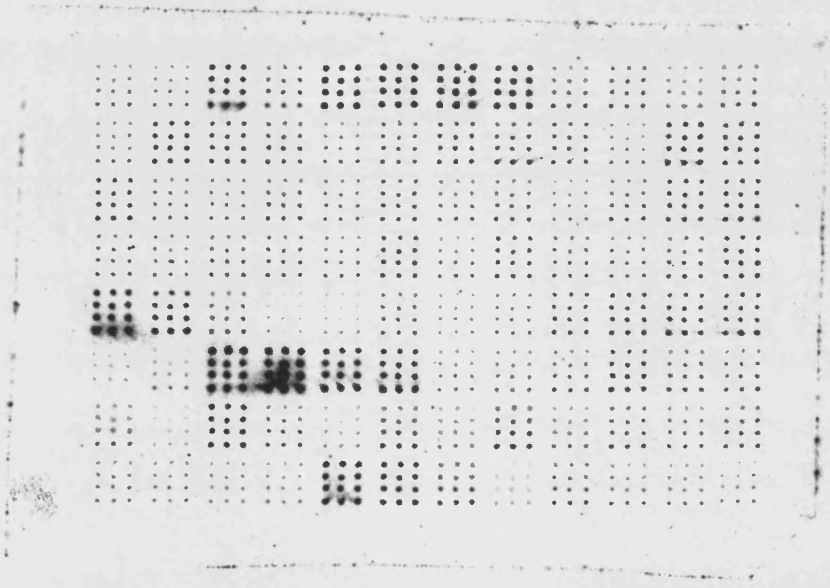
The results of this hybridisation are shown in Figure 4.14. The non-specific binding seen previously with target derived from uninfected cells was no longer present. Instead, only the housekeeping gene probes were seen to bind target on the control hybridisation (as seen in Figure 4.14A). With target derived from infected cells, almost all probes were seen to bind target, and there was also a large range of signal strengths (as seen in B). These results suggested using pin-spotted arrays resulted in much reduced levels of non-specific binding, compared to hybridisations with vacuum-spotted arrays.

Figure 4.14 (following page). Hybridisation of Biomek pin-spotted arrays.
These arrays featured each probe at 4 concentrations: 100ng/µl, 50ng/µl, 25ng/µl and 12.5ng/µl, with the most concentrated towards the bottom of each set. Each probe was spotted as one of three replicates, resulting in sets of 12 spots per probe. A. Control hybridisation representing RNA from uninfected cells. Signals for housekeeping genes can be seen towards the bottom of the array. B. Hybridisation representing NIH 3T3 cells infected with MHV-68 for 12h. The layout of these arrays is as shown in Appendix II. The bottom right corner of each array was cut to aid orientation.

A – 0h pi



B – 12h pi



4.4.10.a Effects of Probe Concentration on Signal Strength

As a range of probe concentrations were spotted onto the Biomek arrays, the effects of changing probe concentration were also observed. The graph in Figure 4.15 shows signal strengths for a probe plotted against the concentration of that probe. Four probes were selected to represent a broad range of signal strengths, from the array shown in Figure 4.14B. All probes showed a linear relationship between 25ng/µl and 100ng/µl. At the lowest concentration of probe, 12.5ng/µl, the linear relationship was lost for probes 1 & 4. As these probes represented higher signals, it seems likely that at the lowest concentration of 12.5ng/µl, these probes were saturated with target. Therefore, 50ng/µl was set as the working probe concentration for future arrays.

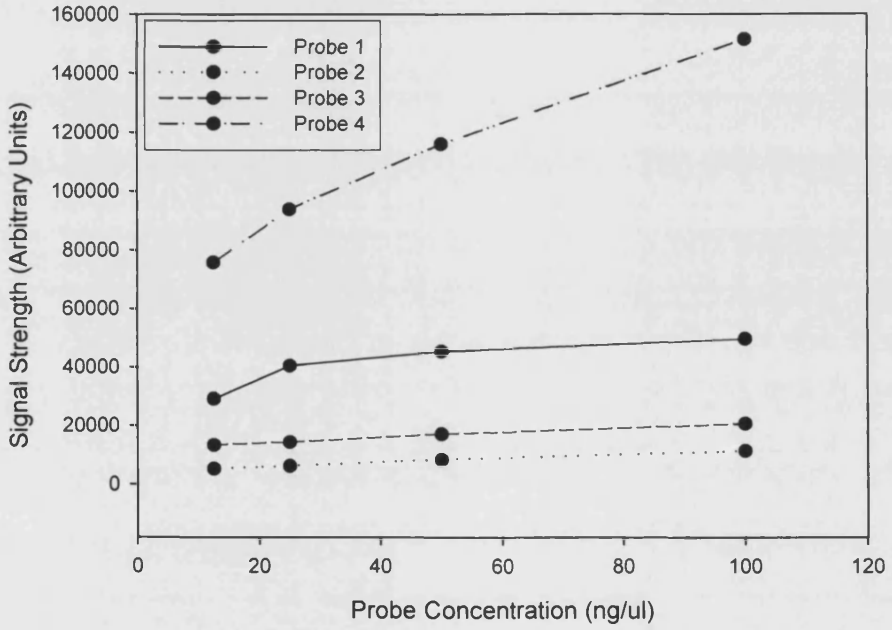


Figure 4.15. The effects of probe concentration on signal strength on DNA arrays.
Four probes were selected to represent the range of signals seen on the array and plotted to show their signal strengths varying with their concentrations.

4.4.10.b Effects of Using Oligo-dT to Prime Target Synthesis

The MHV-68 array system had been designed to use a mix of probe-specific primers to prime the target labelling reverse transcription. However, a comparison between specific priming and oligo-dT priming was also investigated here to clarify the differences resulting from these alternate methods. The hybridisations of the target resulting from these 2 priming methods are shown in Figure 4.16. Generally,

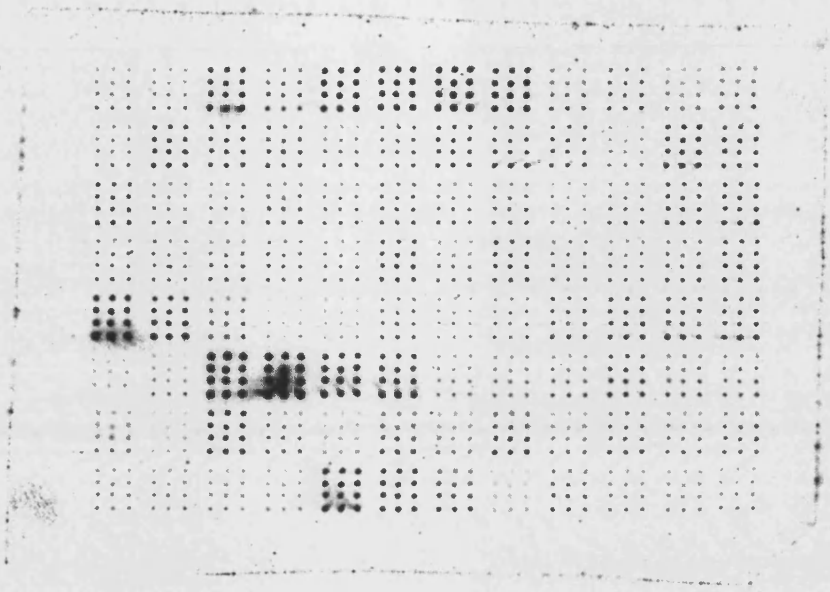
the signals seen on the array hybridised to target synthesised via a specific primer mix (Figure 4.16A), are stronger than those seen when oligo-dT was used (Figure 4.16B). It is possible that the labelled DNA targets produced via specific priming are shorter and therefore bind more efficiently to the probes on the array, compared to those produced by oligo d(T) priming. Furthermore, the relative signal strengths observed after hybridisation differed slightly depending on the priming method. This is also likely to be due to the nature of the targets that result from the priming method used. Specific primers will tend to produce approximately 300bp products, whereas oligo d(T) primed products will be the length of the RNA transcripts.

The results of the hybridisation suggested that the specific primer mix was the more efficient method of labelling target, compared to oligo-dT primers. This validated the design of the array, and therefore, the specific array primer mix was retained as the method for priming RNA.

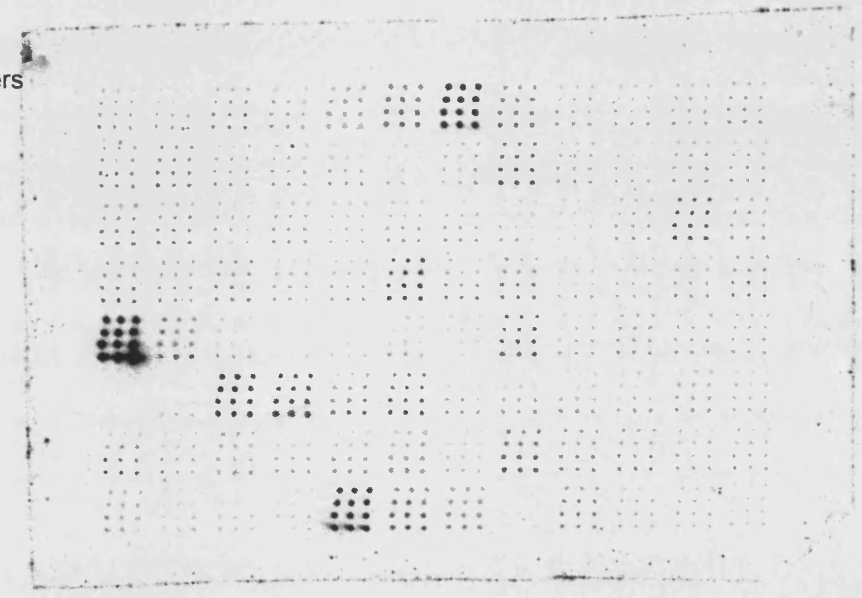
Figure 4.16 (following page). Comparison of 2 priming methods for DNA array target synthesis.

RNA was isolated from NIH 3T3 cells infected for 12h by MHV-68. Target was prepared by reverse transcription of the RNA using A. a specific primer mix, or B. oligo-dT primers. The results of hybridising the targets to arrays are shown above. The layout of these arrays is as shown in Appendix II. The bottom right corner of each array was cut to aid orientation.

A. Specific primers



B. Oligo-dT primers



4.4.11 Manual Pin-Spotted Arrays

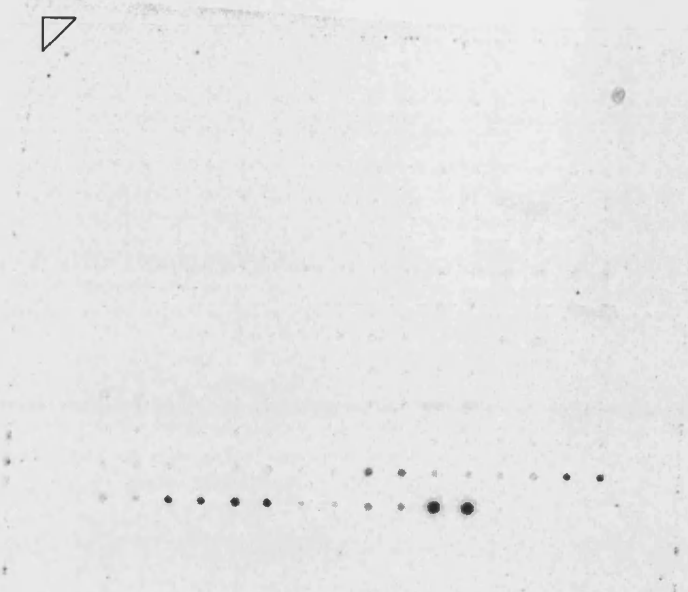
Unfortunately, the Biomek robot became unavailable for use and therefore a manual pin-spotting tool was adopted (384-pin multiblotter; V&P Scientific). This tool has the same droplet volume as the tool used by the robot, and therefore the same protocol for producing arrays was used as before.

On this array, 2 replicates of each probe were spotted at a concentration of 50ng/µl. The layout of the manual pin-tool arrays are shown in Appendix II. The top left corner of each array was cut to aid orientation.

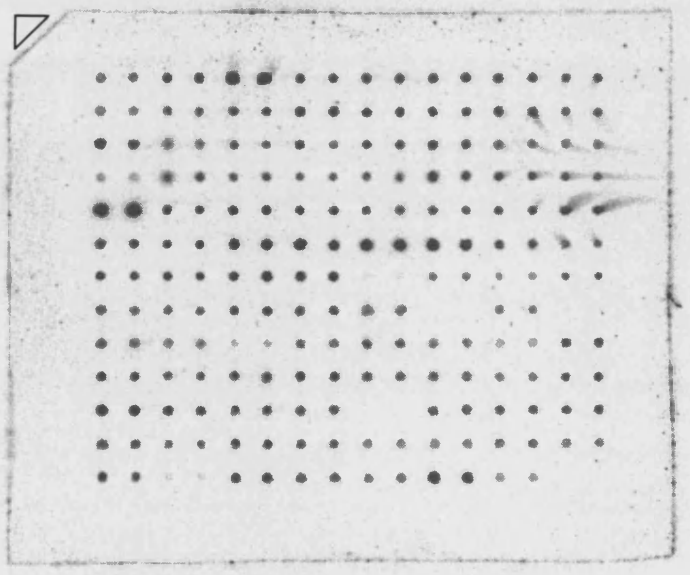
An example of a hybridisation with a manually pin-tool spotted array is shown in Figure 4.17. Mock-infected and 12h pi sample hybridisations are shown. The housekeeping gene probes found towards the bottom of the arrays are seen to bind target in both hybridisations. Viral probes are seen to bind target only in the 12 pi sample hybridisation.

Figure 4.17 (following page). Examples of hybridisation to pin-tool spotted DNA arrays. NIH 3T3 cells were A. mock-infected, or B. infected for 12h by MHV-68. RNA was isolated and target synthesised, before hybridisation. To aid orientation, the top left corner of each array was cut, and is denoted above by a triangle. The layout of the arrays is as shown in Appendix II.

A. Mock-infected



B. 12h pi



4.5 RNA Template Titration

Target synthesis, array manufacture and hybridisation conditions had now been optimised. However, the amount of starting RNA template had yet to be optimised. The reverse transcription of target cDNA had been optimised using 5µg total RNA as template. A series of hybridisations were now performed to investigate the effects that varying RNA template concentration had on the signal strengths seen on the array.

The hybridised arrays were scanned on a phosphoimager and the signals for each probe quantified. The mean signal for each probe was then plotted on a bar chart, as shown in Figure 4.18. The 3 panels show 5µg, 10µg or 20µg of template, respectively. All 3 graphs were plotted on the same scale to help aid comparison of varying RNA template concentration. A smaller graph set into each panel shows the results of each hybridisation plotted on a scale appropriate to that hybridisation. These smaller graphs allow comparison of the profile for each array. In other words, the larger graphs allow a quantitative comparison of hybridisations (due to the absolute scale), while the smaller graphs allow a qualitative comparison (due to the relative scale).

As RNA template concentration was increased, the strength of signal also increased. This was particularly evident for the stronger signals, e.g. F3, F4, F5, and F6. As signal strength relates to the amount of target binding to probe, this indicated that more RNA resulted in more labelled target cDNA.

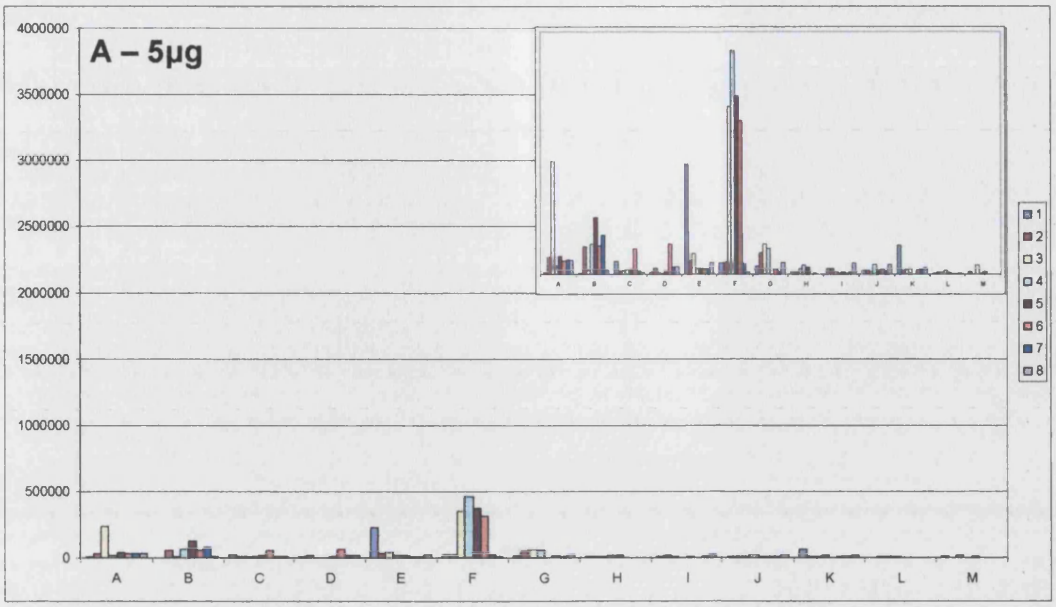
While the stronger signals were seen to get increasingly stronger with RNA concentration, the weaker signals (e.g. B1-8) did not increase significantly in intensity between 10µg (Figure 4.18b) and 20µg (Figure 4.18c) of template RNA. This suggested that there was some selectivity during the labelling reaction for those RNA species present in more abundant quantities.

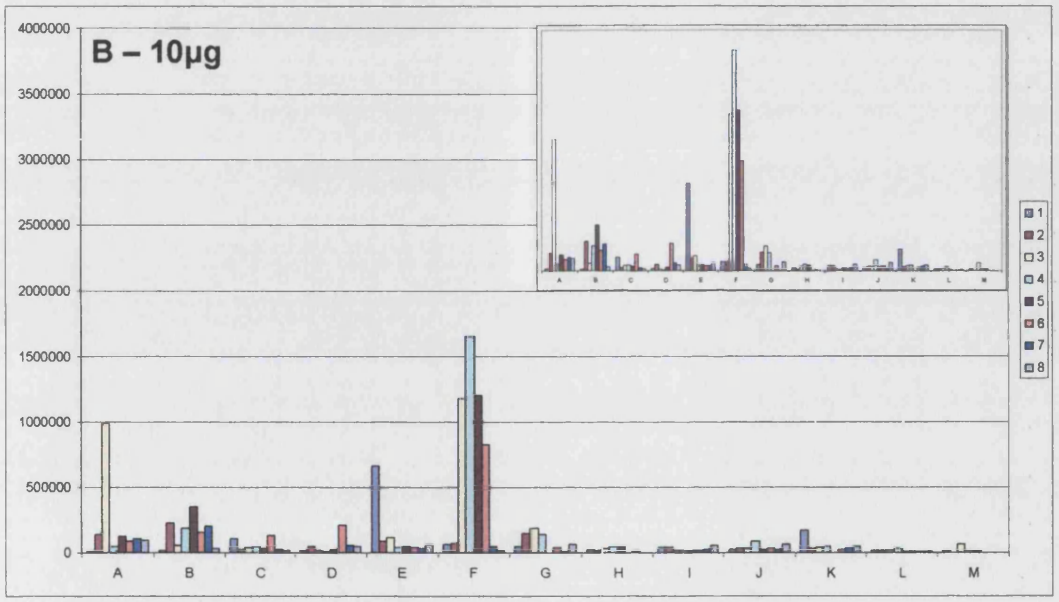
Plotting each array's signals on a relative scale for each array, as shown in the smaller graphs, made the observations above clearer. Whilst the overall profiles for hybridisations at each RNA template concentration were very similar, as RNA concentration was increased, the profile appeared to be stretched along the *y*-axis. This resulted in the difference between the largest and smallest signals being accentuated when 20µg of RNA was used (Figure 4.18c). This also resulted in the difference between smaller signals becoming smaller, making differentiation of these small signals more difficult.

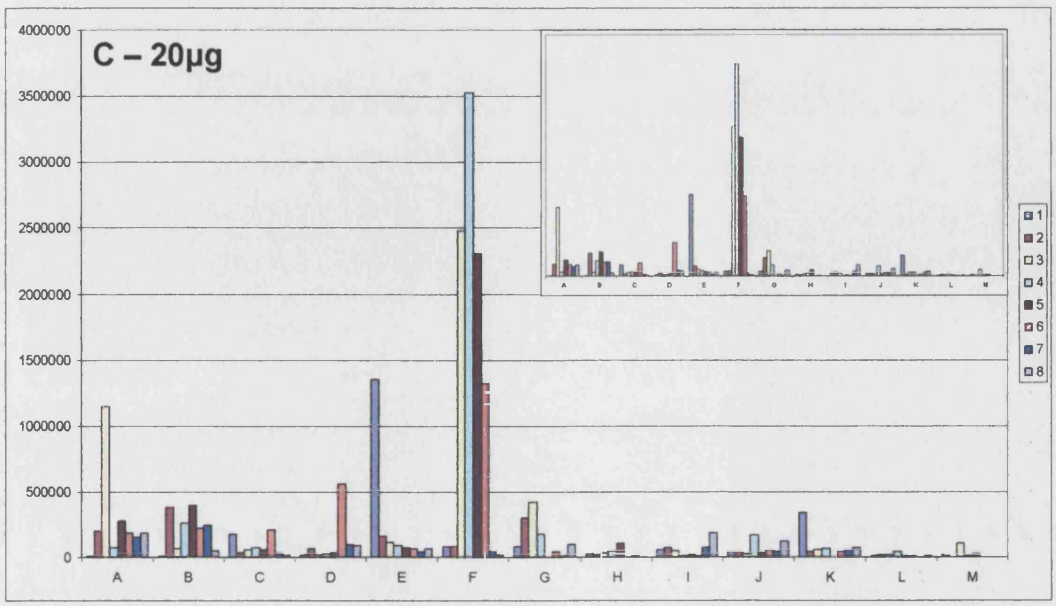
High signal strengths were desirable as were signals that could be easily differentiated from each other, at any level of signal intensity. Therefore, 10µg was set as the amount of RNA template to be used for future arrays.

Figure 4.18 (following page). Effects of varying template RNA concentration on final hybridisation signal strengths.

RNA isolated from NIH 3T3 cells infected for 24h was isolated and used as template to produce labelled target that was hybridised to arrays. The signals on the array were quantified on a phosphoimager and plotted for each probe. Each probe is identified by a letter and number shown on the x-axis and in the key. The signal strength is shown on the y-axis in arbitrary units. A set of smaller graphs are presented within each panel, to show the overall profile for each RNA concentration relative to itself. Each probe was designated a letter and number and corresponds to a gene, as listed in Table 4.1. A. The results of using 5µg total RNA template. B. The results of using 10µg total RNA template. C. The results of using 20µg total RNA template. The layout of these arrays is as shown in Appendix II.


A – 5µg
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
A B C D E F G H I J K L M
1 2 3 4 5 6 7 8


B – 10µg
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
A B C D E F G H I J K L M
1 2 3 4 5 6 7 8


C – 20µg
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
A B C D E F G H I J K L M
1 2 3 4 5 6 7 8

4.6 Normalisation of Array Data

While optimising the array, it had been noted that the housekeeping gene probes consistently showed very low signals when analysing infected samples. A comparison between the housekeeping gene signals from arrays hybridised with target derived from uninfected cells, and target derived from infected cells is shown in Figure 4.19. It is clear that MHV-68 infected cells show a significant reduction in levels of housekeeping gene RNA, relative to uninfected cells. Therefore, the transcription of housekeeping genes could not be used for normalising data from separate arrays, as their levels were shown to vary through the course of an infection.

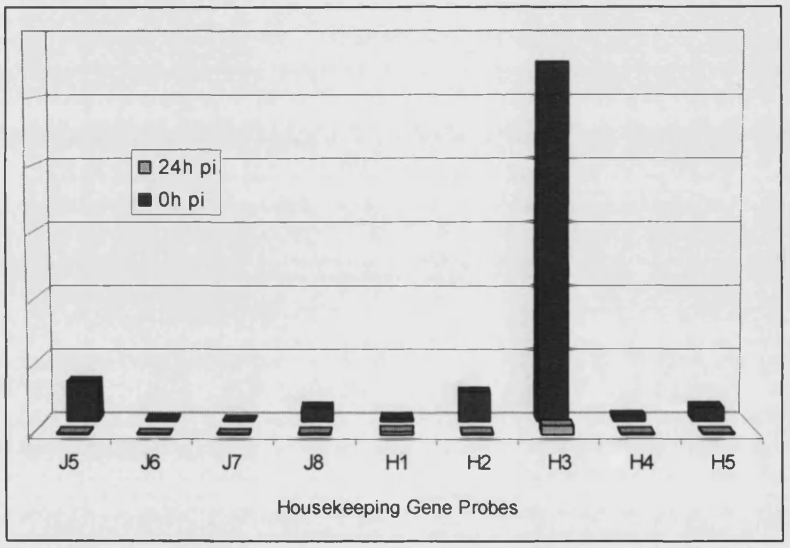


Figure 4.19. Comparison of signals for housekeeping gene probes between arrays hybridised to target derived from RNA of uninfected cells and those infected for 24h. Housekeeping gene probe signals for the uninfected cell array are plotted behind, with the corresponding signals from the infected array plotted to the front of the graph.

As a method of normalisation for data from separate arrays was required, a separate internal control was developed. An additional probe to luciferase mRNA was designed, and incorporated into the MHV-68 array. To act as an internal control, 10ng of luciferase RNA was added to each RNA sample, prior to synthesis

of labelled cDNA target. All hybridisation signals on arrays could now be expressed as a signal ratio relative to the luciferase signal. As the hybridisation signal for luciferase corresponded to the same quantity of RNA, signals from separate arrays (that would be subject to variations in labelling efficiencies, hybridisation efficiencies etc) could be compared directly.

To assess the efficiency of this process of normalisation, NIH 3T3 cells were infected with MHV-68 for 5h and then analysed with the array. This was then repeated. Each step from initial infection to quantification of signal was performed independently for each array analysis. Background was calculated as the mean of signals recorded for the negative control probes for TMV and plasmid (but not the "water probe", which was excluded from background calculations). The background subtracted signals from these 2 independent array analyses were then plotted against each other. Linear regression returned a line of gradient 0.80, and the 2 datasets had a correlation coefficient of r = 0.76 (see Figure 4.20A). This showed that data derived from independent array experiments did not match each other quantitatively. The data was now normalised by conversion into signal ratios relative to the internal luciferase control. The same graph as in panel A was re-plotted using normalised data. Linear regression of this plot returned a line of gradient 0.99, while the correlation coefficient for the 2 datasets was r = 0.94 (see Figure 4.20B). This showed that normalisation of data from independent arrays was successful by this method. Therefore, the luciferase normalisation would be used before comparing data from all subsequent arrays.

Finally, to simulate comparing multiple arrays with each other, the same process was repeated a further 2 times. This allowed a cross-comparison of 4 independent analyses of transcription at 5h pi. The scatter plot is shown in Figure 4.20, panel C. Linear regression returned a gradient of 1.01 and a correlation coefficient of r = 0.91. This confirmed that multiple arrays could be compared directly to each other using luciferase normalised data. It also served to further reinforce confidence in this method of normalisation.

Furthermore this experiment was a measure of the reproducibility of the MHV-68 array. The gradient of 1.01 suggested that the array was indeed a reproducible system, based on 4 independent array experiments.

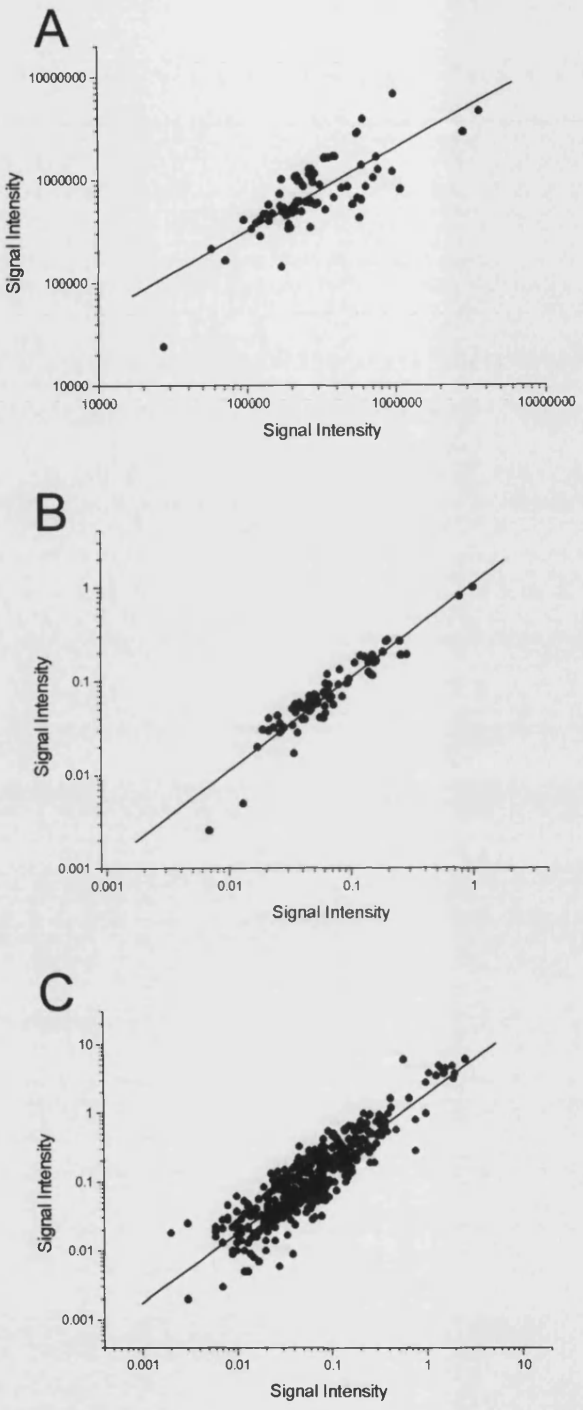


Figure 4.20. Scatter plots of DNA array data showing effects of normalisation.

A. Background-subtracted data from two repeated arrays (time point 5h pi) were plotted against each other. Linear regression analysis gives a line of gradient 0.80 (r = 0.76). B. The data from the array in (A) were subjected to normalisation via the internal control and these normalised data were plotted against each other. Linear regression analysis gives a line of gradient 0.99 (r = 0.94). (C) Reproducibility of arrays was assessed by plotting four repeated arrays (time point 5h pi) against each other in all combinations. Linear regression analysis gives a line of gradient 1.01 (r = 0.91).

4.7 Assessing Sensitivity of Arrays

Having optimised the array system and a method to allow comparison of data across arrays, its sensitivity was now assessed. NIH 3T3 cells were infected with MHV-68 at moi's ranging from 1pfu/10^6 cells – 20pfu/cell, and RNA isolated 5h pi. The RNA was used as template to reverse transcribe labelled target, which was hybridised to arrays.

Mean signal values were calculated from the 2 replicates of each probe, and normalised to the internal luciferase control. The global transcription profiles at 5h pi, following infection at moi 1 - 0.001 are shown in Figure 4.21A-D. The strength of transcription signals were seen to decrease with moi. The detection of a representative selection of transcripts, ORF 52-58, is shown in panel E of Figure 4.21. The selection includes both abundantly and weakly expressed genes. This graph shows the results for infections at moi 1pfu/cell – 1pfu/10^6 cells.

These results suggest a lower limit of sensitivity for the MHV-68 array, of 1 infected cell in 100 cells. However, a working limit of 1 infected cell in 10 was set to ensure that all experimental data were significantly within the limits of sensitivity.

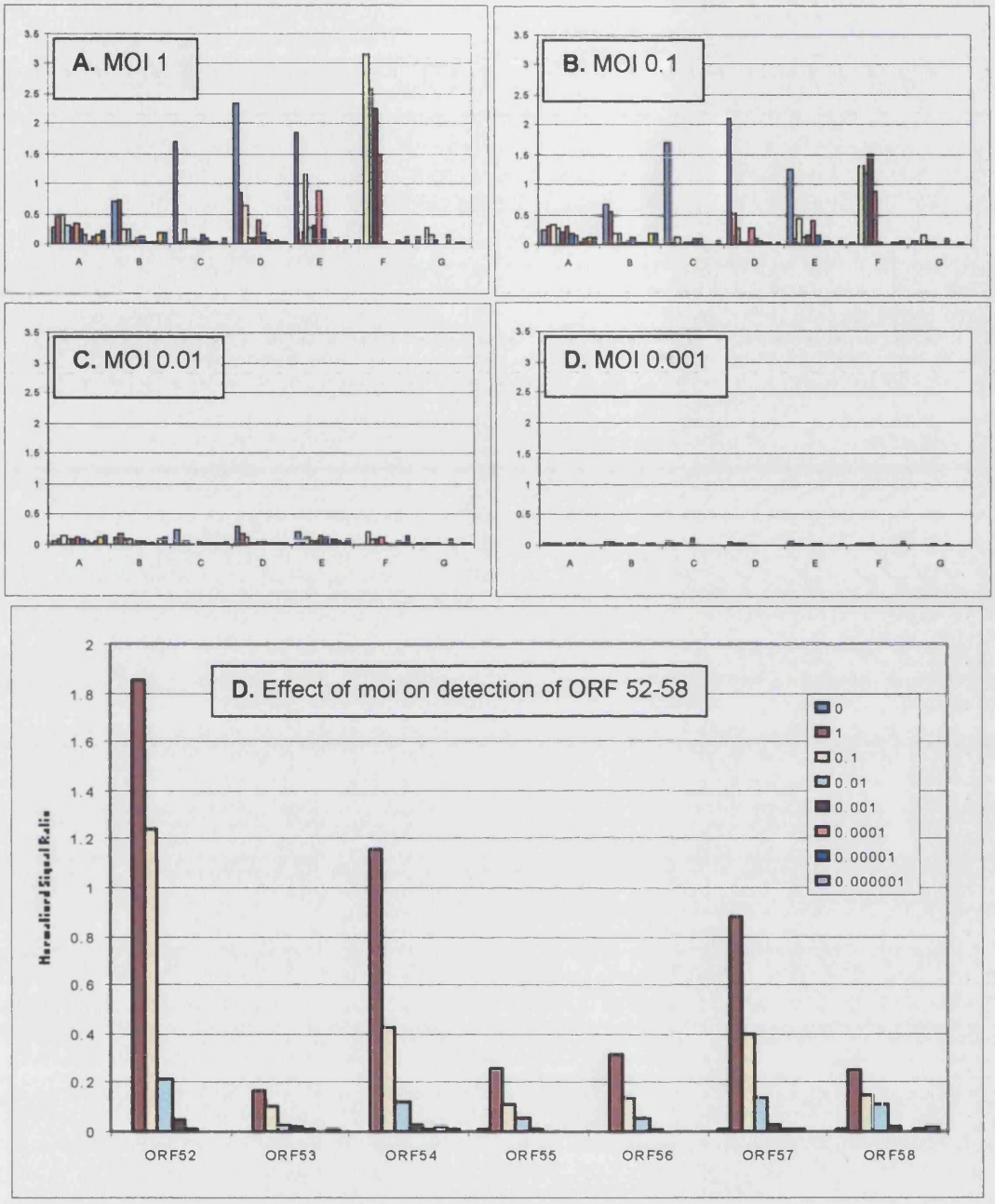


Figure 4.21. Sensitivity of the MHV-68 array.

NIH 3T3 cells were infected with MHV-68 at a range of moi's. RNA was isolated 5h pi and used as template for reverse transcription of labelled target. Following hybridisation, the normalised signal ratios were calculated and are shown here as bar charts. Global transcription profiles are shown in panels A – D, representing infections at moi 1 (A), 0.1 (B), 0.01 (C) and 0.001 (D) pfu/cell. E. A representative selection of transcripts is shown separately (ORF 52-58). The selection was made to include abundantly and weakly expressed transcripts. Moi's are indicated as a value of pfu/cell. The layout of these arrays is as shown in Appendix II but the probe signals are shown in a different order to that in Figure 4.18.

4.8 Discussion

The aim of the work presented in this chapter was to design and develop an array system to study the transcriptional profile of MHV-68 at a global level. Similar array systems already in use were used as templates and guides for this array system (Zhu, 1998; Chambers, 1999; Stingley, 2000; Jenner, 2001). A membrane-based cDNA probe array system hybridising radio-labelled cDNA target was decided upon as a suitable system for this analysis of MHV-68 gene expression. This solution provided the best combination of performance and practicality.

Bioinformatics software was used to design primer sequences for the amplification of probe sequences using MHV-68 DNA as template. Every primer was designed to have the same biochemical properties. This allowed the whole set of 3' primers to be used together to prime the labelling reaction. Software was also used to minimise the possibility of cross hybridisation between primers, and non-specific binding between probes and targets.

Each probe sequence was cloned and sequence verified. Attempts to sequence probes G5 and G8 were unsuccessful, possibly due to their high GC content. However, as they appeared to be the correct probe sequence based on PCR amplification with probe-specific primers, they were included on the array as non-sequence verified probes. A high efficiency protocol for amplification of probes from plasmid stocks was developed to allow simple generation of probe stocks for spotting onto arrays.

The actual array itself was based on a charged nylon membrane. Probe sequences were spotted onto these membranes using a manual 384-pin tool. Probe concentrations of 25ng/µl to 100ng/µl were found to show a linear relationship with signal strength. The concentration of probe for spotting onto array membranes was set at 50ng/µl.

Total RNA was used as the initial template for the MHV-68 array. RNA was isolated from samples using the optimised RNA isolation protocol developed in Chapter 3.2. RNA template concentration was set at 10µg following hybridisations with varying amounts of template. These hybridisations showed that increasing RNA template increased signal strengths overall, which made reading the arrays clearer and any non-specific background levels were smaller in comparison. However, with increased template, the highest signals became higher and the lowest signals became lower, relative to each other. Ten µg total RNA was set as the template concentration.

The labelling reaction was found to return consistently better results when using Superscript II reverse transcriptase. This enzyme also made it possible to increase incubation temperature and time which greatly increased yield. Primer concentrations were increased to 0.2mM. The hybridisation process was optimised to reduce non-specific binding by adding C_0T-1 DNA to the hybridisation solution.

An overview of the optimised protocol is shown in Figure 4.22. The specifics of the optimised protocol are as follows:

1. Isolate total RNA from samples using the protocol developed for the differential display system, including DNAse treatment, quantification and quality assessment.
2. 10μg RNA was spiked with 10ng of luciferase mRNA, and primed with the array primer mix (0.2μM) for 5min at 70°C, then place on ice.
3. Add the reverse transcription reaction mix: 1x first strand buffer, 1mM dT/G/CTP, 2.5μCi/μl α-^{33}P dATP and 5U/μl Superscript II reverse transcriptase, in a final volume of 20μl. Incubate for 90min at 50°C. The reaction was terminated by adding 1μl 100mM EDTA, before placing on ice.
4. The labelled target was purified using the Nucleospin kit and denatured at 95°C for 5min prior to hybridisation.
5. At the same time, arrays were denatured by placing them on a filter paper soaked with denaturation buffer (0.66M NaCl, 0.5M NaOH) for 10 min and then washing in ddH_2O for 10min. Finally they were washed in 40mM phosphate buffer pH7.3 for 10min.
6. Arrays were prehybridised at 60°C for at least 30min in Expresshyb containing salmon sperm DNA and murine C_0T-1 DNA.
7. The denatured target was hybridised to arrays in fresh Expresshyb with salmon sperm DNA and murine C_0T-1 DNA at 60°C for approximately 16h.
8. The arrays were washed 3 times for 15min each at low stringency. They were then washed another 3 times for 15min each at high stringency.
9. Arrays were covered in cling film and used to expose phosphor screens.
10. Arrays signals were quantified using a phosphoimager. The data was background subtracted and normalised to the luciferase signal. Analyses of the data were performed.

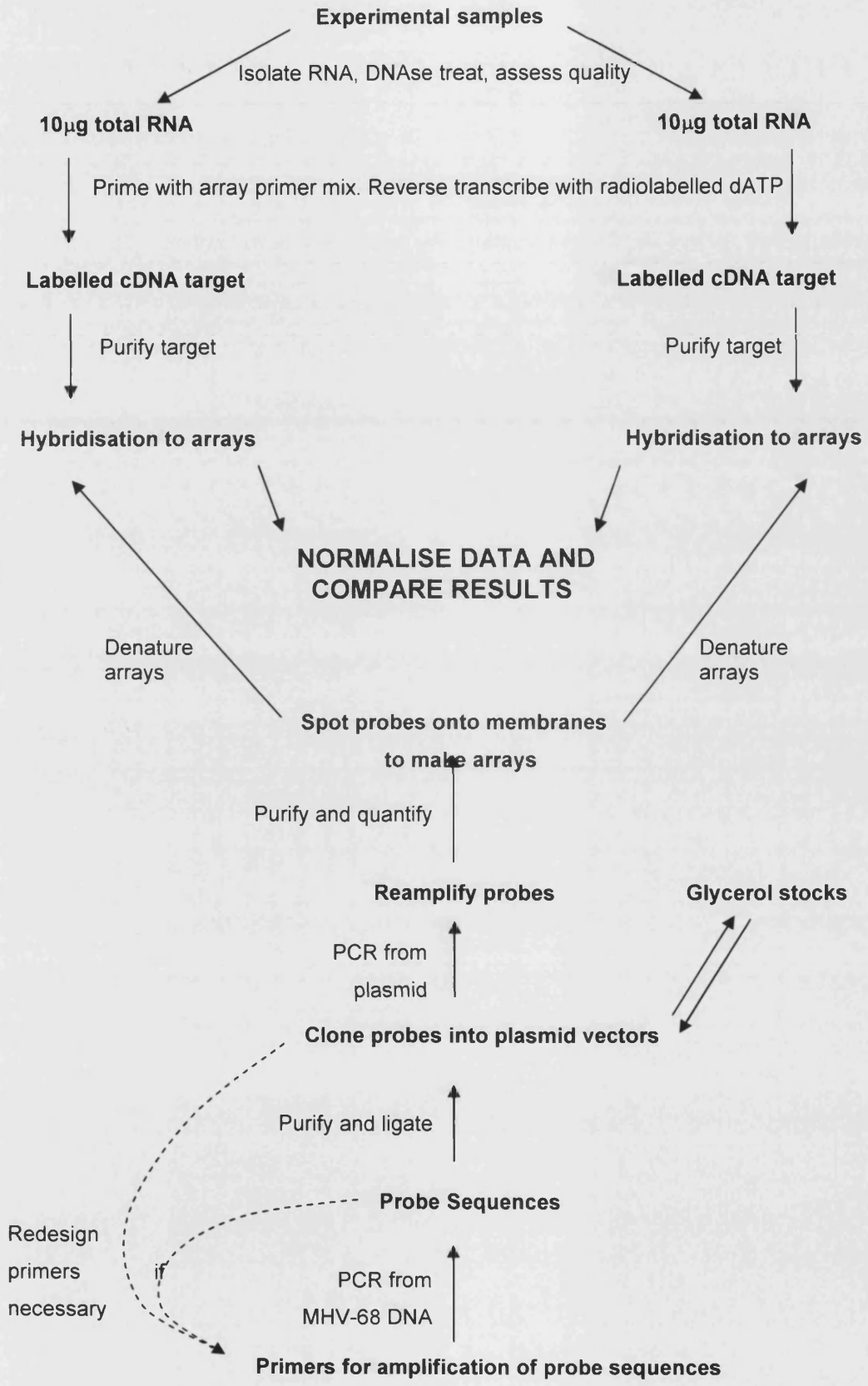


Figure 4.22. Overview of the MHV-68 Array system.

The optimisation of the MHV-68 array was performed bearing consideration to the global nature of the system. It was important that the system was capable of distinguishing each of the individual probe signals that make up the global transcription profile. Such a system should be able to process data across a large dynamic scale, encompassing the potential transcription levels of each gene. In this way the data produced by the array data becomes more than the sum of its individual probes, as it also provides a complete picture of gene expression, one gene's expression relative to all other genes represented on the array.

However, developing a working array system proved to be complicated. The series of optimisation experiments showed that the methods used to produce arrays were important factors in the successful subsequent application of arrays. The vacuum-spotted arrays appeared to bind high levels of background signal and therefore did not produce reliable data. The pin-spotted arrays however, were successfully used. The main difference between these 2 array systems was the nature of the probes. The probes on the vacuum-spotted arrays were much larger in diameter and covered a larger area of the membrane than their counterparts on the pin-spotted arrays. While the total amount of DNA in each probe on the 2 arrays differed greatly, the concentration of DNA per unit area was equal for both arrays. The much higher quantities of probe DNA on the vacuum-spotted array could result in increased background to signal ratios, as the total background noise would increase with the amount of probe DNA, whereas actual signal would only increase until all target DNA was hybridised.

Another key difference between the two arrays was that the probe DNA was denatured *in situ* for the pin-spotted arrays as opposed to the vacuum-spotted array probes, which were denatured prior to spotting. It is possible that the probes on the vacuum-spotted arrays were not fully denatured by the time the probes were fixed onto the array membrane, which would reduce the specificity of the vacuum-spotted arrays. This would also have the result of increasing the background to signal ratio. It seems likely that these differences between the 2 arrays were responsible for the difference in results, when the 2 arrays were tested.

With a working MHV-68 array, other factors such as hybridisation conditions and the target synthesis reaction could be optimised. Sensitivity and reproducibility of the MHV-68 array system was then assessed and a method of normalisation developed to allow cross-comparison of array data.

This process from design through optimisation to a final working array system has now provided the opportunity to investigate MHV-68 at the level of gene expression.

The array technology will now be implemented to perform this investigation. However, it remains necessary to test the array system in a real experimental situation, before its efficacy can be fully determined.

5 DNA Array Analysis of MHV-68 Transcription during an *In Vitro* Lytic Infection

While there have been several gene expression studies of gammaherpesvirus latency (Heller, 1982; Hall, 2000) and also reactivation (Miller, 1994; Zalani, 1996; Speck, 1997), few have focused on *de novo* productive infection due to a lack of suitable *in vitro* systems in which to study it. The aim of the work presented here was to perform a global study of MHV-68 gene expression using the membrane-based DNA array system, designed and optimised as described in Chapter 4. To this end, a productive lytic infection was followed in a permissive tissue culture system.

5.1 DNA Array Analysis of MHV-68 Transcription In Vitro

Confluent NIH 3T3 cell monolayers were infected with extracellular virus at moi 10 and cells harvested at 1, 3, 5, 8, 12 and 18h pi. Mock-infected cells were also taken at 1h and 24h pi. Total RNA was extracted at each time point, reverse transcribed and the resulting radiolabelled cDNA targets hybridised to arrays, using the optimised protocols described in Chapter 4.

The signals from each array were quantified and mean signal ratios were calculated for each probe from the replicate arrays (n = 4, except 1h pi (n = 2) and 5h pi (n=6)). The data for each time point was plotted as a bar chart and can be seen in Figure 5.1 - Figure 5.7. This series of hybridisations showed that MHV-68's global transcriptional profile changes as it progresses through its life cycle, and that individual genes have individual expression kinetics.

5.1.1 DNA Array Analysis of Uninfected Cells

The MHV-68 array includes probes for 9 housekeeping genes. When mock-infected NIH 3T3 cells were screened with the array, a range of signals for these housekeeping genes were observed (Figure 5.1). Of these, ubiquitin was the most highly expressed, showing approximately half the signal observed for the luciferase probe. Signals for GAPDH, β-actin, ribosomal protein S29, and HPRT were evident as well. The probes for myosin 1, MOD, Cab45 and phospholipase showed very little signal. This suggests that there were only low levels of these transcripts in confluent NIH 3T3 cells, and that the abundance of these transcripts were below the sensitivity threshold of the array. It may be possible to optimise the sequence of the probes for these genes and increase their sensitivity. The housekeeping genes were selected based on general maintained expression in a number of

physiological states (4.2.3), and therefore it was not surprising that some were not highly expressed in NIH 3T3 cells. As expected, the 100ng of luciferase RNA added to the template RNA produced one of the highest signals on the array, both here and on subsequent infected sample arrays.

There were also a few viral probes (ORF 35, 46, M10c-72, M14) that bound very low levels of target. Although levels of non-specific hybridisation were estimated and subtracted from the array data, this shows that it was not completely effective as no viral RNAs were present in this RNA sample. It may be possible to optimise the background probes to more accurately simulate non-specific hybridisation to the array probeset, and also to make the viral probes more specific to their target cDNA. For the purposes of this analysis, an arbitrary cut off point of 0.05 was set. This arbitrary value was twice the highest background signal seen on the array hybridised to mock-infected samples.

RNA samples was also isolated from mock-infected cells 18h pi and analysed on the array. This hybridisation showed the same results as those from mock-infected samples taken 1h pi (data not shown). This indicated that any changes in housekeeping gene transcript levels, observed between 1h and 18h pi, were likely to be due to the effects of MHV-68 infection.

5.1.2 DNA Array Analysis of Cells Infected for 1h

The viral transcription profile at 1h pi (Figure 5.2) was similar to that seen in mock-infected cells (shown in Figure 5.1) as almost no viral signals were observed. Again, a small number of viral probes (ORF 4, 35 46, M14) bound low levels of target, but none were above the cut-off point (5.1.1). However, even at this early time after infection, the abundance of housekeeping gene transcripts was reduced relative to their levels in uninfected samples. This suggested that MHV-68 very quickly down-regulated host transcription, within 1h pi. Rapid shutoff of host gene expression has also been observed in other herpesvirus infections (Read, 1983; Bigger, 2002).

Fig. 5.1. Transcript Abundances in Mock-Infected Cells

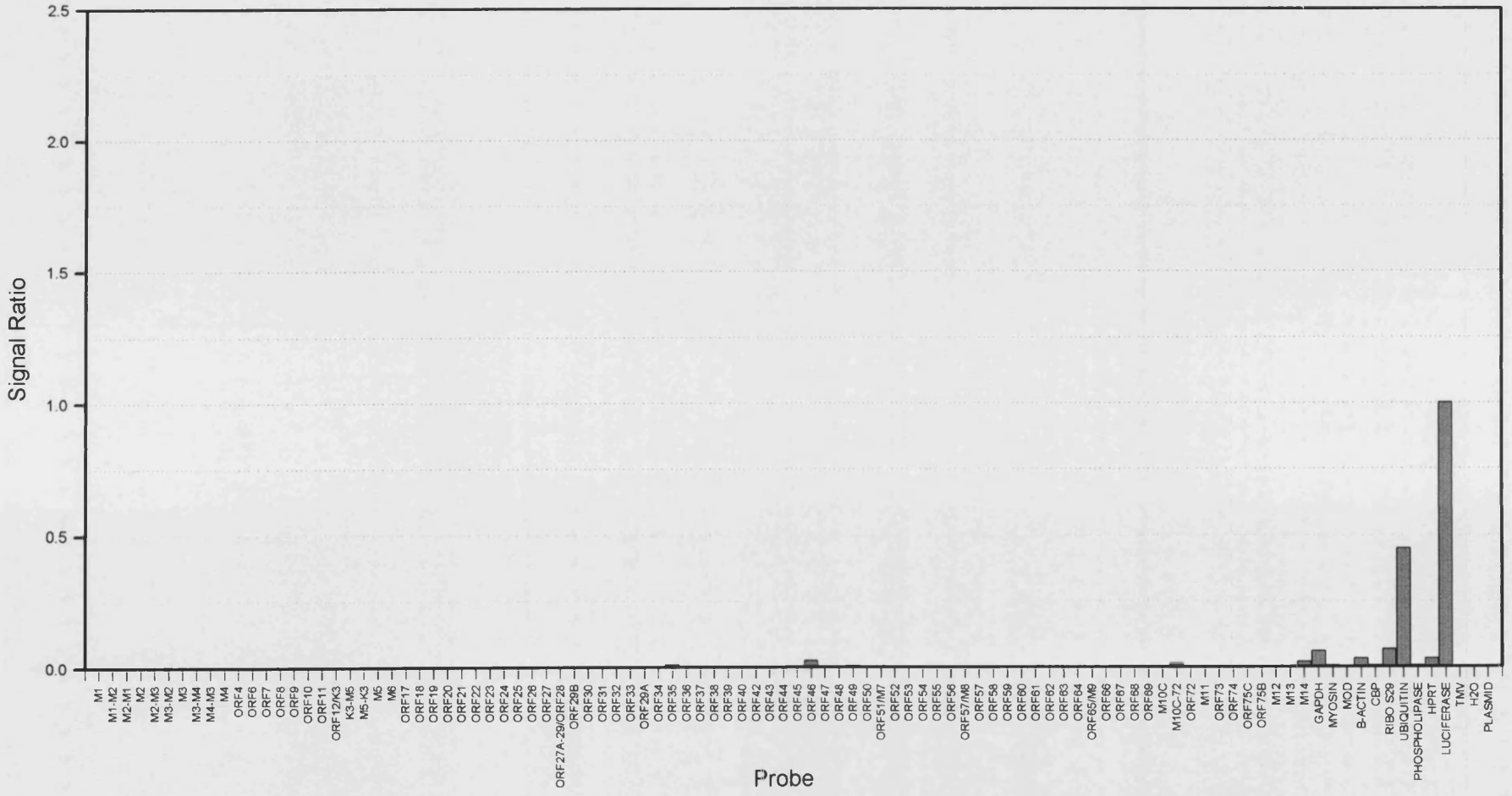


Fig. 5.2. Transcript Abundances at 1h pi

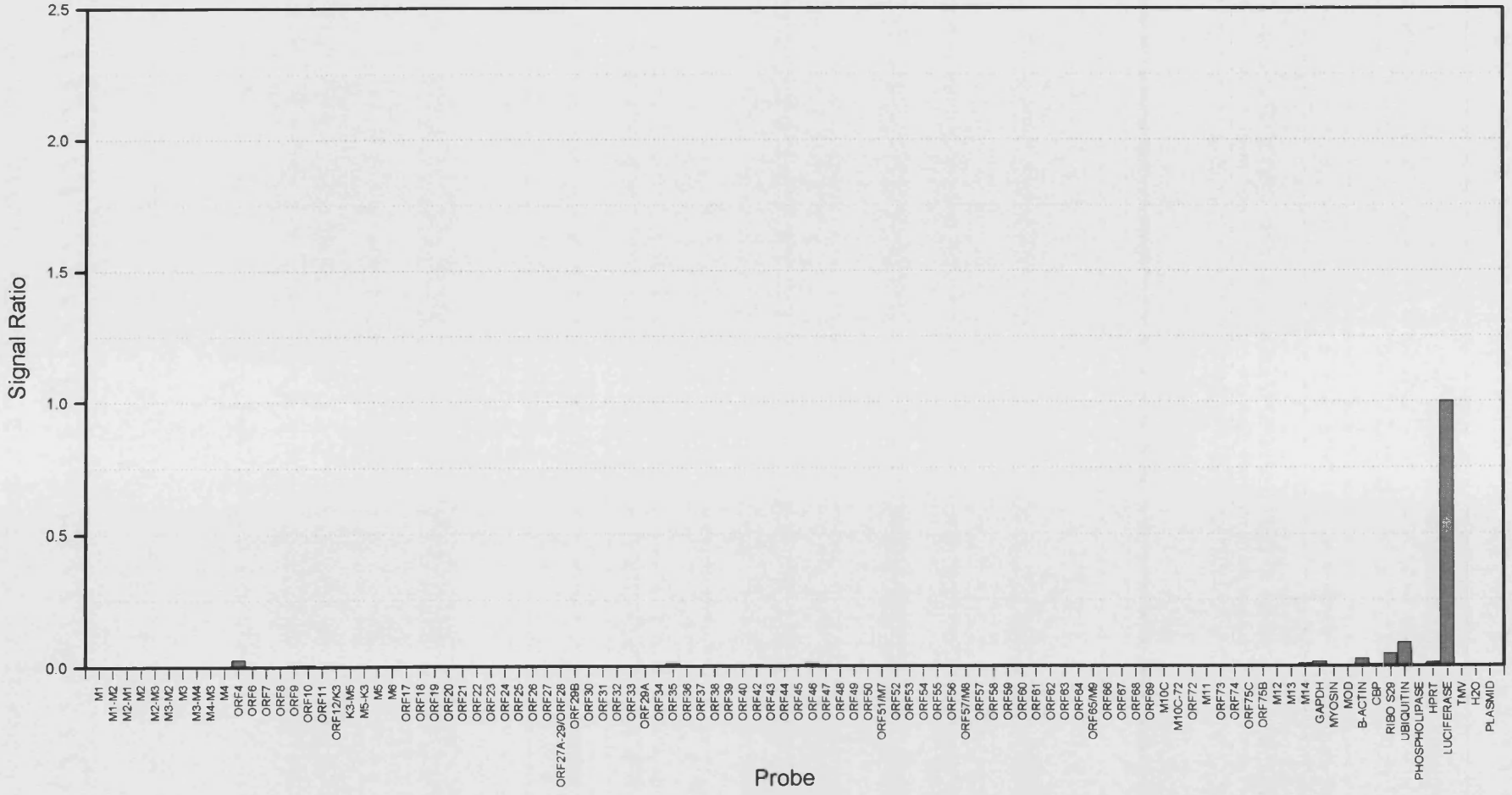


Fig. 5.3. Transcript Abundances at 3h pi

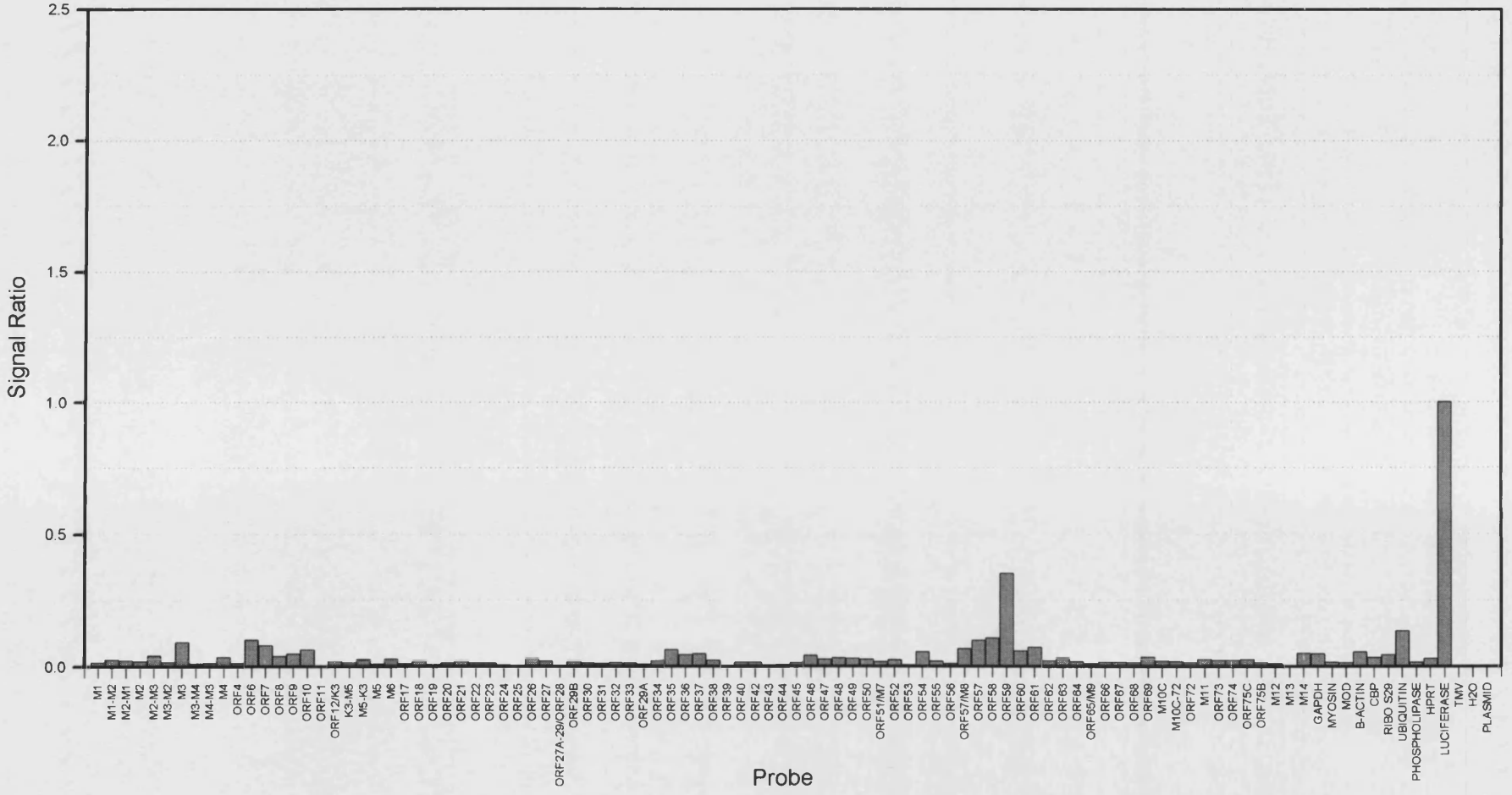


Fig. 5.4. Transcript Abundances at 5h pi

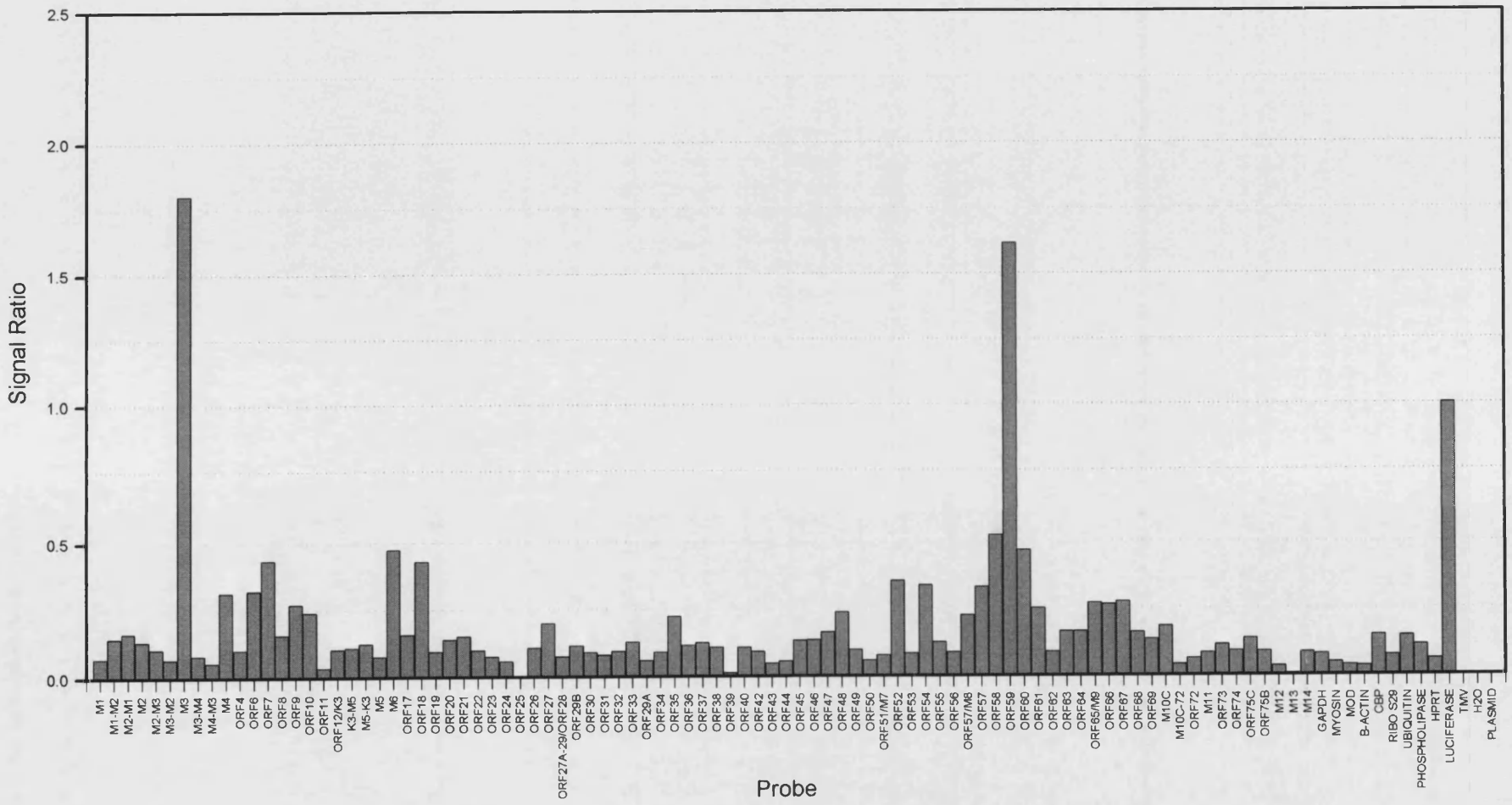


Fig. 5.5A. Transcript Abundances at 8h pi

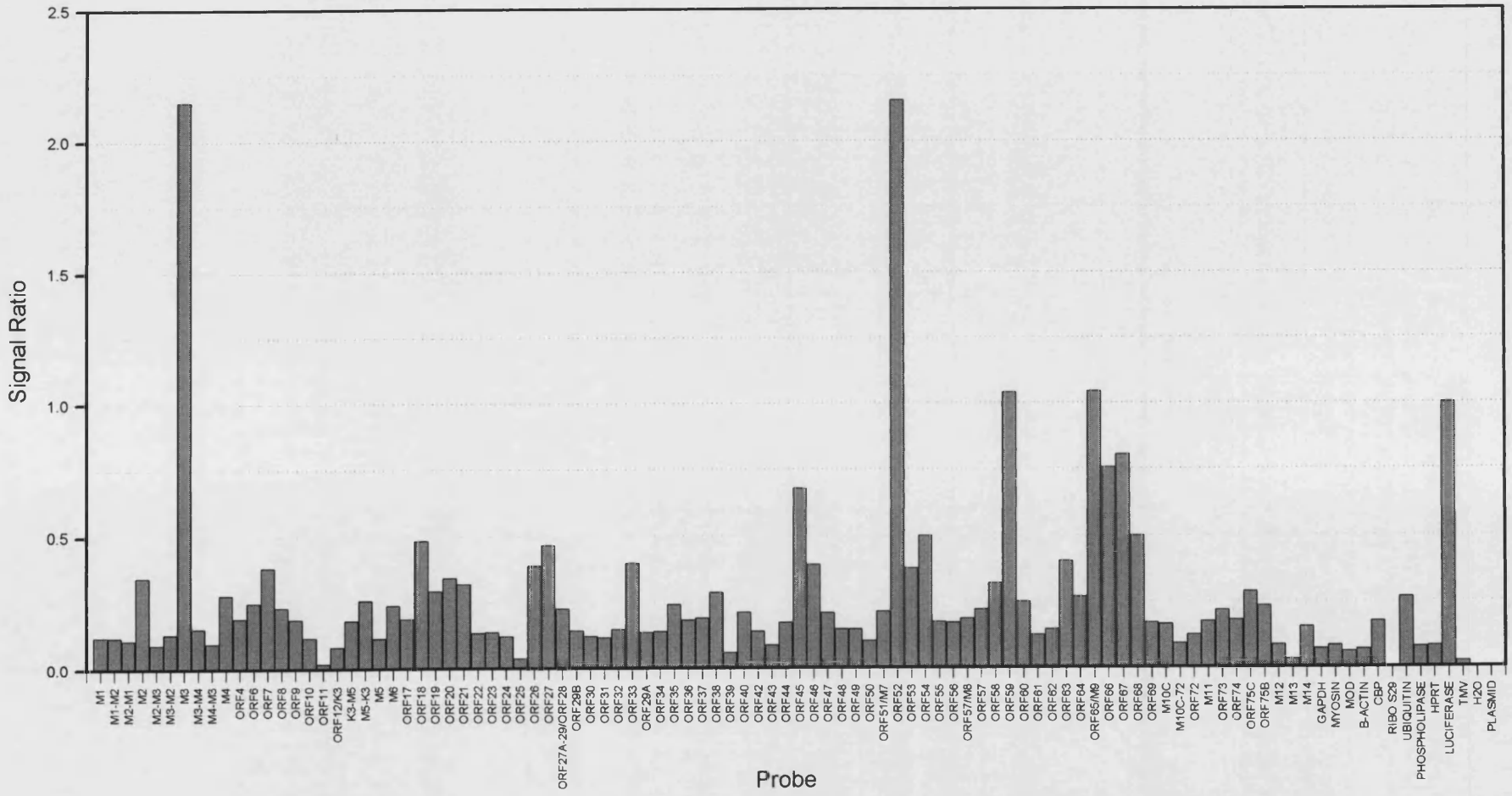


Fig. 5.5B. Changes in transcript abundances between 5h and 8h pi

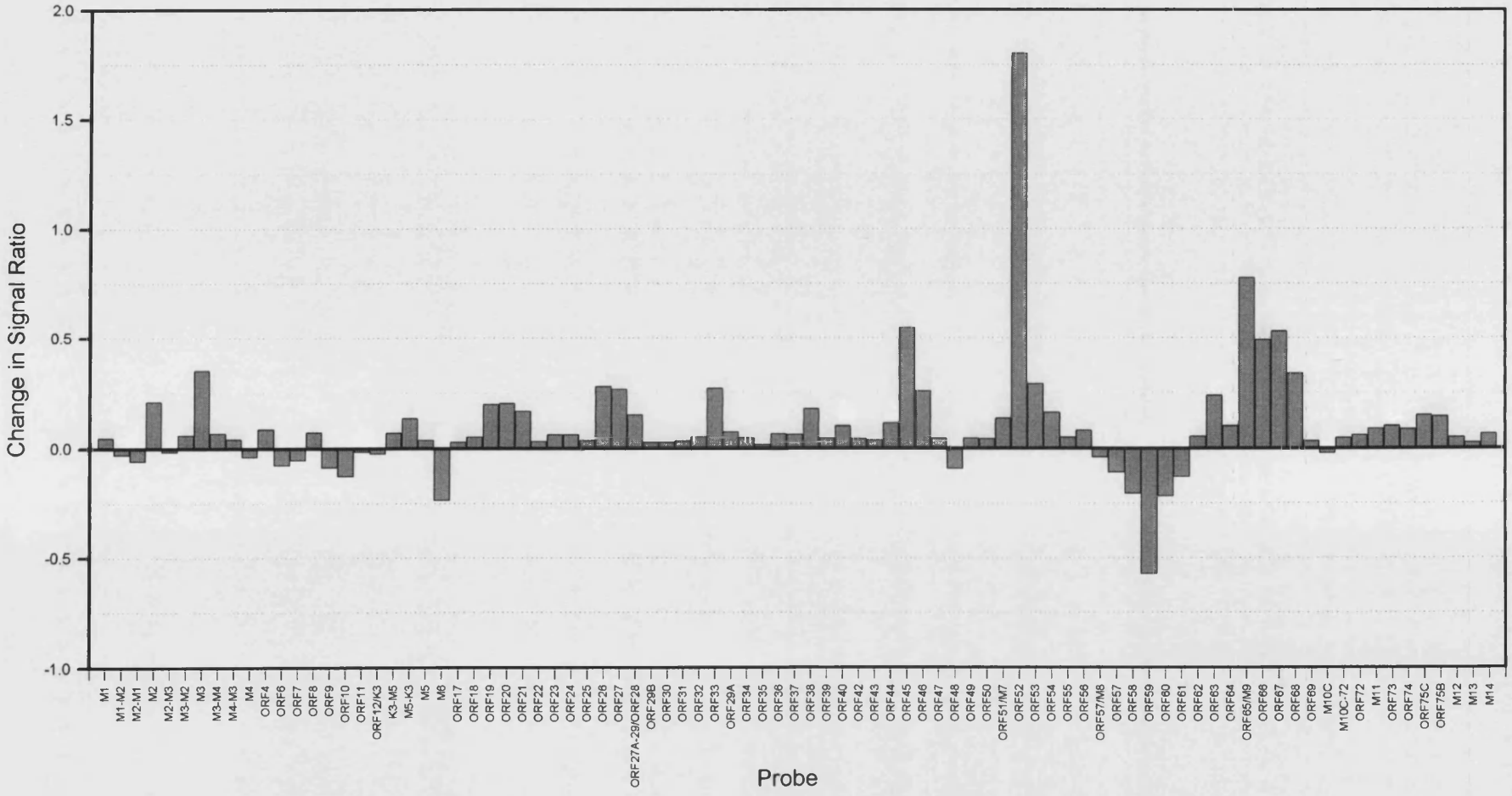


Fig. 5.6A. Transcript Abundances 12h pi

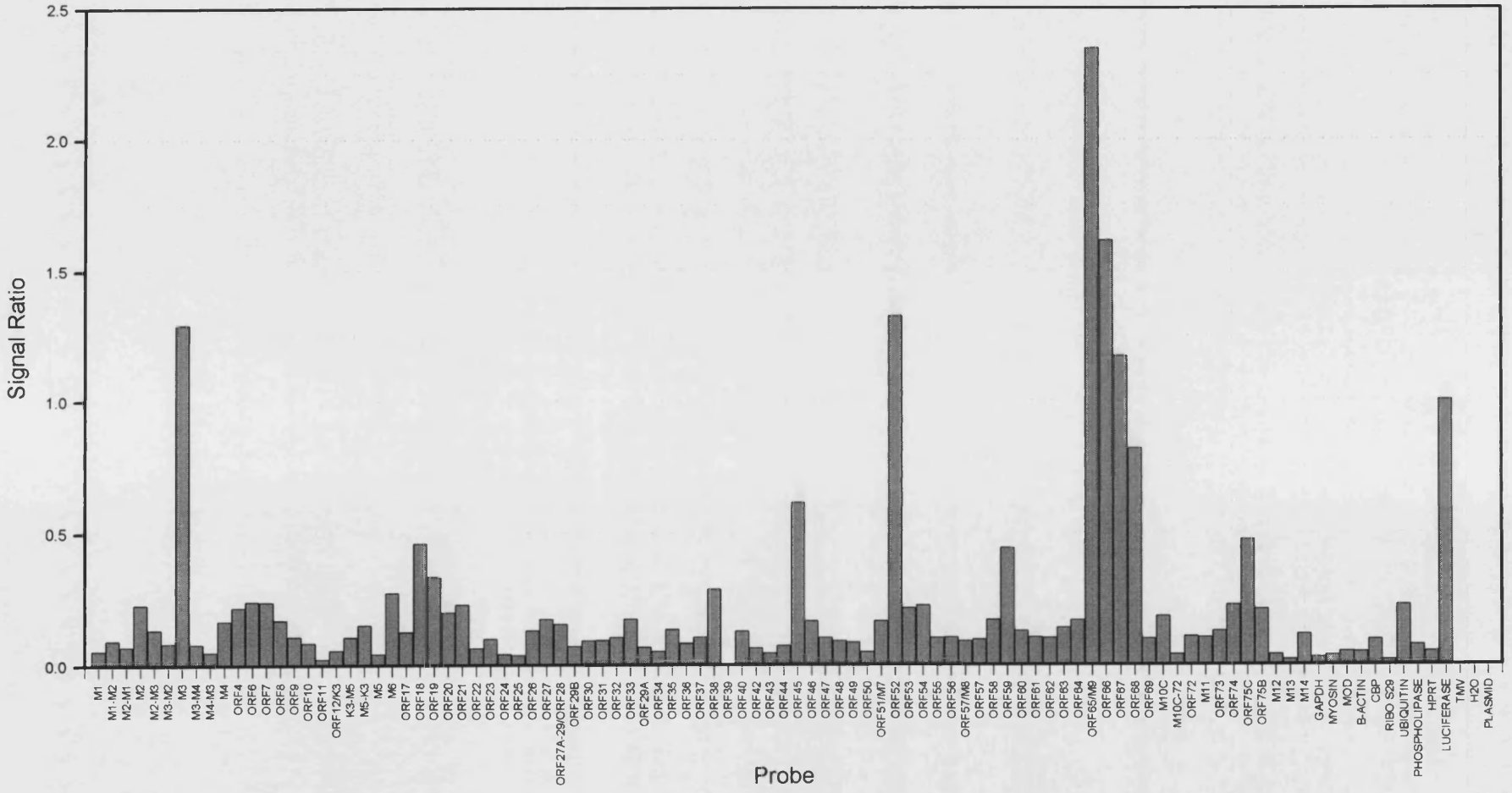


Fig. 5.6B. Changes in transcript abundances between 8h and 12h pi

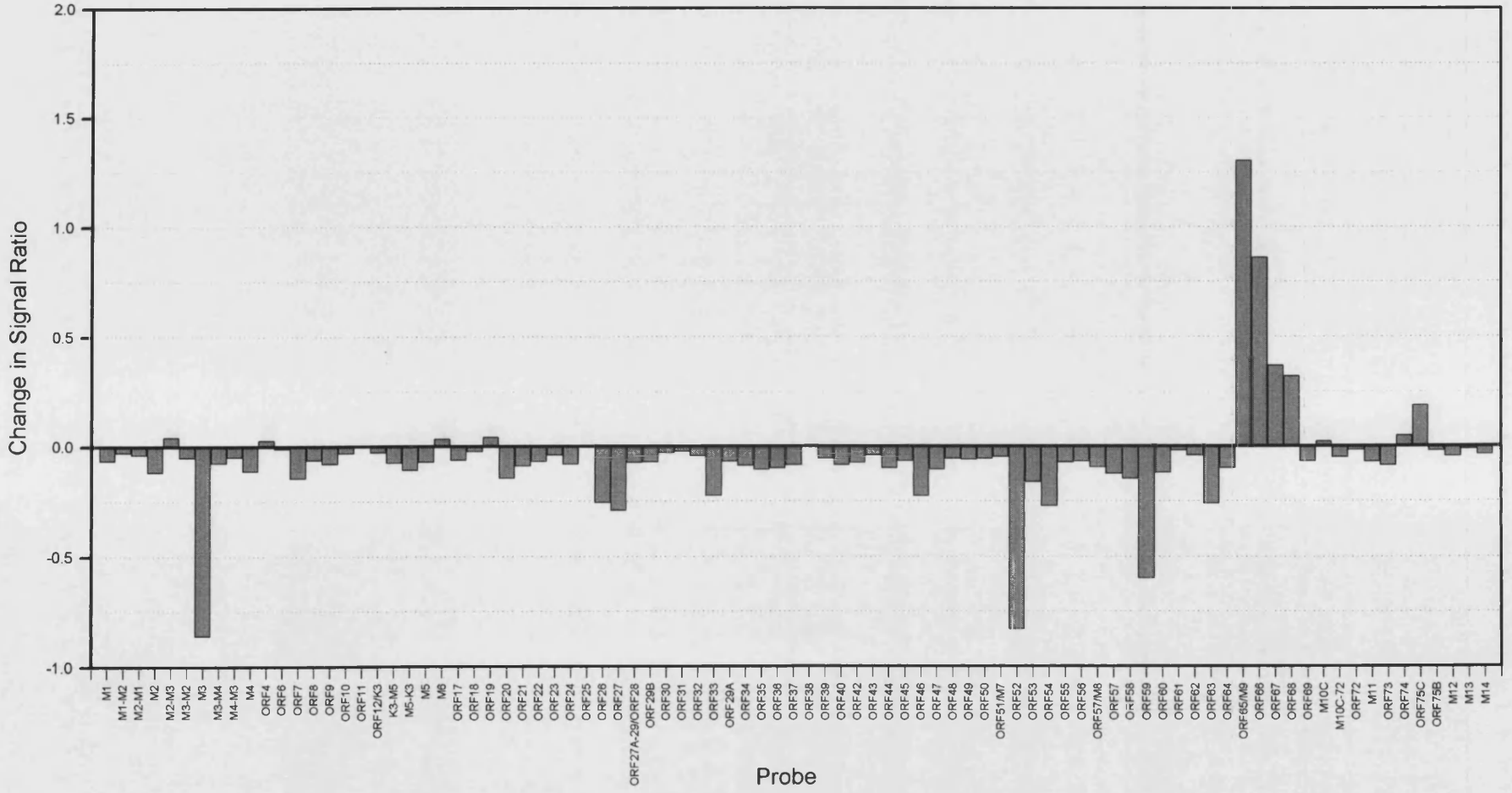


Fig. 5.7A. Transcript abundances at 18h pi

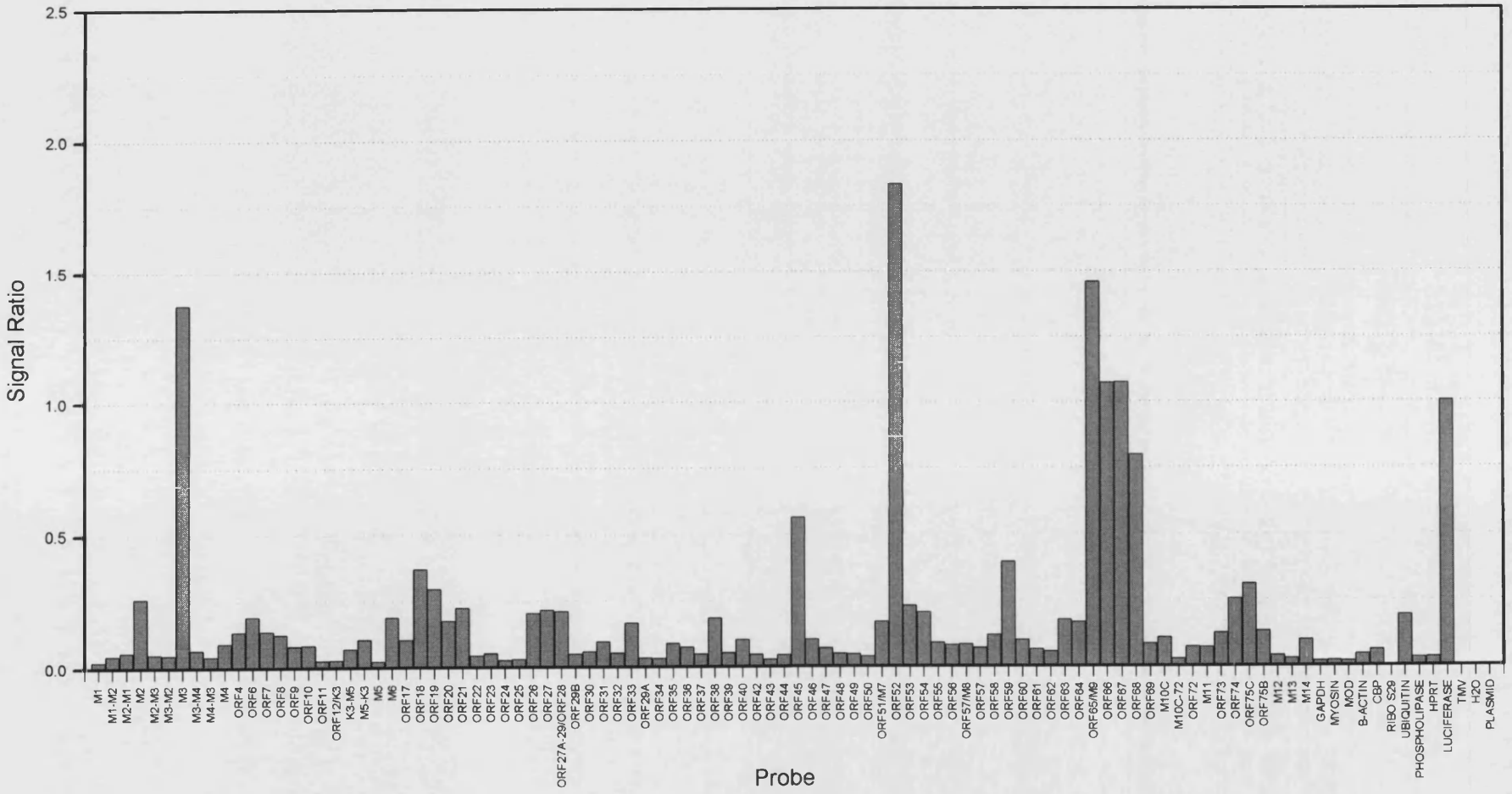


Fig. 5.7B. Changes in transcript abundances between 12h and 18h pi

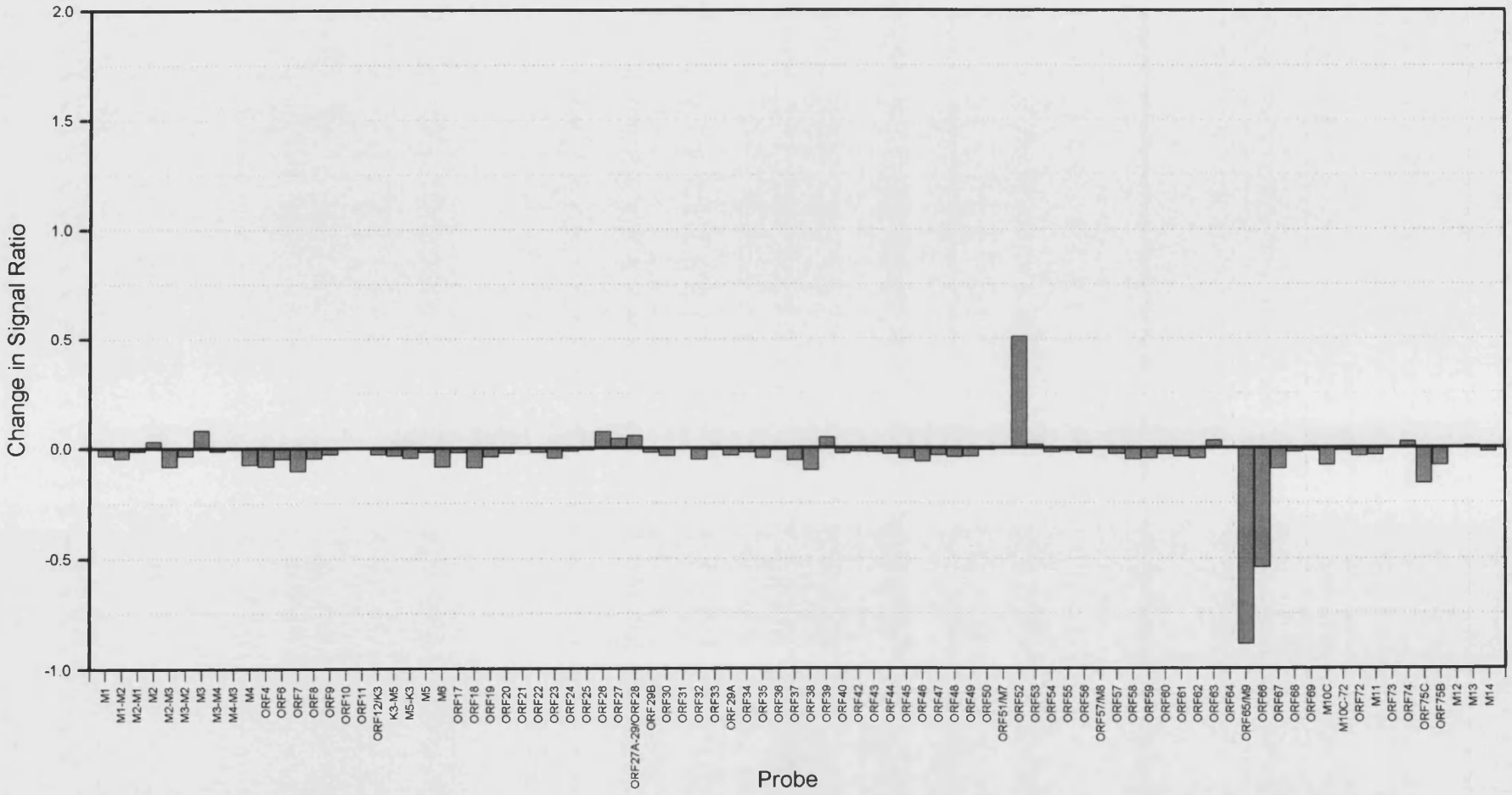


Figure 5.1. DNA array analysis of mock-infected cells.
NIH 3T3 cells were mock-infected and harvested and RNA isolated for the synthesis of radiolabelled cDNA target. Target was hybridised to the MHV-68 DNA array, and signals quantified with a phosphoimager. The mean signals for each probe were normalised against the luciferase signal. The dataset was then plotted to aid analysis. The bar chart shown represents the mean dataset of independent experiments (n = 4).

Figure 5.2. DNA array analysis of cells infected by MHV-68 1h pi.
NIH 3T3 cells were infected and harvested 1h pi. The array hybridisation was performed and resulting data analysed as detailed for Figure 5.1. The bar chart shown represents the mean dataset of independent experiments (n = 2).

Figure 5.3. DNA array analysis of cells infected by MHV-68 3h pi.
NIH 3T3 cells were infected and harvested 3h pi. The array hybridisation was performed and resulting data analysed as detailed for Figure 5.1. The bar chart shown represents the mean dataset of independent experiments (n = 4).

Figure 5.4. DNA array analysis of cells infected by MHV-68 5h pi.
NIH 3T3 cells were infected and harvested 5h pi. The array hybridisation was performed and resulting data analysed as detailed for Figure 5.1. The bar chart shown represents the mean dataset of independent experiments (n = 6).

Figure 5.5. DNA array analysis of cells infected by MHV-68 8h pi.
A. NIH 3T3 cells were infected and harvested 8h pi. The array hybridisation was performed and resulting data analysed as detailed for Figure 5.1. The bar chart shown represents the mean dataset of independent experiments (n = 4). **B.** Changes in transcript abundances between 5h and 8h pi. The relative changes in signal ratios, between the 2 time points for each probe were plotted. The relative change was calculated by subtracting the signal ratio at the previous time point, from the current signal ratio. An increase in signal ratio between the 2 time points is represented as a bar rising above the x-axis, and a decrease as a bar falling below the x-axis.

Figure 5.6. DNA array analysis of cells infected by MHV-68 12h pi.
A. NIH 3T3 cells were infected and harvested 12h pi. The array hybridisation was performed and resulting data analysed as detailed for Figure 5.1. The bar chart shown represents the mean dataset of independent experiments (n = 4). B. Changes in transcript abundances

between 8h and 12h pi. The relative changes in signal ratios were calculated as detailed for Figure 5.5.

Figure 5.7. DNA array analysis of cells infected by MHV-68 18h pi.
A. NIH 3T3 cells were infected and harvested 18h pi. The array hybridisation was performed and resulting data analysed as detailed for Figure 5.1. The bar chart shown represents the mean dataset of independent experiments (n = 4). **B.** Changes in transcript abundances between 12h and 18h pi. The relative changes in signal ratios were calculated as detailed for Figure 5.5.

5.1.3 DNA Array Analysis of Cells Infected for 3h

At 3h pi there was still little viral transcription overall (Figure 5.3) but the signals for some transcripts were above the cut-off point. They are listed in Table 5.1. Of these, ORF 59 (homologous to the processivity subunit of herpesvirus DNA polymerase) was detected at much higher levels than any other transcript, with a signal ratio of 0.35. Interestingly, most of the genes listed in Table 5.1 are predicted to encode proteins involved in DNA replication (ORF 6, 7, 54, 59, 60, and 61). Similarly, in HSV-1, the ribonucleotide reductase genes are transcribed at a very early stage of infection (Swain, 1986; Wymer, 1992; Hanson, 1994).

In addition to DNA replication, other highlighted transcripts consisted of a chemokine-binding protein (ORF M3), a predicted α protein (ORF 57) and also, 2 structural genes (ORF 35, 58) and a DNA packaging protein (ORF 7). M3 is known to be a chemokine binding protein (Parry, 2000) and it is interesting to see that a gene involved in immune system evasion is expressed at this early time point post-infection. α gene transcripts would be expected to be detected at this early stage of infection, and therefore the putative function of α-protein assigned to ORF 57 seems corroborated by its relatively high transcript levels 3h pi. The detection of 2 structural genes and a DNA packaging protein is a little surprising as these γ gene products are not required by the virus at this early stage of infection. However, γ1 genes are transcribed prior to viral DNA replication and this is likely to be the case here. ORF 10 was also highlighted, but there are no functional predictions for this ORF.

Table 5.1. Viral transcripts present at relatively high levels (more than twice the background signal) in NIH 3T3 cell infected for 3h by MHV-68.

Probe	Putative Function
M3	Chemokine binding protein (Parry, 2000)
6	ssDNA binding protein
7	DNA packaging
10	Unknown
35	Tegument protein
54	UTPase
57	α protein
58	Membrane protein
59	DNA polymerase, processivity subunit
60	Ribonucleotide reductase, small subunit
61	Ribonucleotide reductase, large subunit

5.1.4 DNA Array Analysis of Cells Infected for 5h

There was a considerable increase in viral transcript levels at 5h pi compared to previous time points (Figure 5.4). This was evident across the whole of the MHV-68 genome, indicating a rapid escalation of viral gene expression between 3h and 5h pi. Indeed, every viral probe showed higher signal ratios at 5h pi than at 3h pi.

The signals for ORF M3 and 59 were significantly higher than those of other genes, with signal ratios of 1.8 and 1.6, respectively. Seventeen probes returned signal ratios greater than 0.25, and these are listed in Table 5.2. The 6 DNA replication genes highlighted at 3h pi, are also highlighted here in Table 5.2, with the addition of ORF 9 (putative catalytic subunit of DNA polymerase). Four genes predicted to encode structural proteins (ORF 58, M9, 66, 67), and 2 genes involved in immune system evasion (ORF M3, 4) are also listed in Table 5.2 (Virgin, 1997; Milligan, 1998). There are also 4 genes with no putative function, which would be interesting to study further.

Table 5.2. Viral transcripts with signal ratios great than 0.25 in NIH 3T3 cells infected with MHV-68 for 5h.

Probe	Putative Function
M3	Chemokine binding protein (Parry, 2000)
M4	Unknown
6	ssDNA binding protein
7	DNA packaging protein
9	DNA polymerase, catalytic subunit
M6	Unknown
18	Unknown
52	Unknown
54	UTPase
57	α protein
58	Membrane protein
59	DNA polymerase, processivity subunit
60	Ribonucleotide reductase, small subunit
61	Ribonucleotide reductase, large subunit
M9	Capsid protein
66	Capsid protein
67	Tegument protein

5.1.5 DNA Array Analysis of Cells Infected for 8h

Between 5h and 8h pi, the viral transcriptome changed considerably (Figure 5.5A). Of 88 viral probes, 68 showed increased signal ratios relative to levels at 5h pi. Many of the transcripts that had been highlighted for their relatively high signal ratios at previous time points were now observed to have fallen in abundance.

Twenty-seven probes returned signal ratios of greater than 0.25, which was an indication of the generally high level of viral gene expression occurring. The sum signal on the array was highest at this time point, which showed that viral transcription was maximal around 8h pi. Seven probes returned signal ratios of greater than 0.5, and these are listed in Table 5.3.

Table 5.3. Viral transcripts with signal ratios great than 0.5 in NIH 3T3 cells infected with MHV-68 for 8h.

ORF	Putative Function
M3	Chemokine binding protein (Parry, 2000)
45	Unknown but required for efficient replication *in vitro*[1]
52	Unknown
59	DNA polymerase, processivity subunit
M9	Capsid protein
66	Capsid protein
67	Tegument protein

[1]Inhibition of ORF 45 expression leads to reduction of viral replication (Jia, 2003)

ORF M3 and 52 still dominated the transcriptional profile at 8h pi, both with signal ratios of 2.15. ORF 59 was also highly expressed, but was the only gene involved in DNA replication with a signal ratio above 0.5. Three structural transcripts were also abundant (ORF M9, 66, 67), as well as 2 genes of unknown function.

However it should be noted that transcripts present in low abundances can be as significant as transcripts that are abundant. Indeed the abundance of a transcript is only one measure of gene expression. For example, changes in expression patterns may be more functionally significant than absolute levels. Therefore, these relative changes in transcript abundance were plotted, as shown in Figure 5.5B.

This shows that the levels of 19 transcripts fell between 5h and 8h pi, as listed in Table 5.4. Levels of ORF 59 had decreased by the greatest extent, which shows that while ORF 59 was one of the most abundant transcripts at this time point, its levels were actually falling. As ORF 9 levels also fell between these 2 time points, it seems likely that the synthesis of DNA polymerase had peaked by 8h pi. In total, 6 genes involved in DNA replication that had previously been highlighted for their high transcript levels, were all found to have fallen in abundance between 5h and 8h pi (ORF 6, 7, 9, 59, 60, 61). This suggests that the synthesis of proteins for replication of MHV-68 DNA had peaked by 8h pi. However, the relationship between synthesis of DNA replication proteins and viral DNA replication remains unclear.

Table 5.4. Viral transcripts decreasing in abundance between 5h and 8h pi with MHV-68.

Probe	Putative Function	Signal Reduction
M1-M2	Unknown	- 0.18
M2-M1	Unknown	- 0.34
M2-M3	Unknown	- 0.02
M4	Unknown	- 0.03
6	ssDNA binding protein	- 0.07
7	DNA packaging protein	- 0.05
9	DNA polymerase, catalytic subunit	- 0.09
10	Unknown	- 0.12
11	Unknown	- 0.02
K3	Bovine herpesvirus 4 α protein homologue	- 0.02
M6	Unknown	- 0.24
48	Unknown	- 0.09
M8	Exon of ORF 57	- 0.04
57	α protein	- 0.11
58	Membrane protein	- 0.21
59	DNA polymerase, processivity subunit	- 0.57
60	Ribonucleotide reductase, small subunit	- 0.22
61	Ribonucleotide reductase, large subunit	- 0.13
M10c	Unknown	- 0.02

Two putative α gene transcripts (ORF K3, 57) fell in abundance between 5h and 8h pi. This was unsurprising for α gene transcription. However, the fall in abundance of a structural protein transcript (ORF 58) was surprising, as a γ gene would be expected to be transcribed up to the later stages of the infection cycle. This, combined with the relatively high levels of ORF 58 transcripts detected at early time points, seems incompatible with its sequence homology to a membrane protein of other herpesviruses. However, examining the genetic organisation around ORF 58 shows that it lies downstream of ORF 59, with which it is predicted to share a polyadenylation site. Therefore, the ORF 59 transcript would contain the ORF 58 sequence, and the ORF 58 probe would detect both ORF 58 and 59 transcripts. As ORF 58 is a highly expressed β gene, it seems likely that the unexpected expression profile for ORF 59 is due to binding of the ORF 59 transcript to the ORF

58 probe. Indeed ORF 59 transcripts were observed to fall in abundance between 5h and 8h pi. Nine probes for unknown genes and inter-genic regions also showed reduced signals at this time point. These would be interesting to analyse further.

5.1.6 DNA Array Analysis of Cells Infected for 12h

At 12h pi, the global transcription profile is in sharp contrast to that seen at 5h and 8h pi, as shown in Figure 5.6A. In particular a cluster of structural genes (ORF M9, 66, 67 & 68) now dominate the transcriptome, with signal ratios of 2.35, 1.61, 1.17 and 0.82, respectively. These represent 4 of the 7 probes with signal ratios above 0.5, which are listed in Table 5.5. ORF M3 and 2 genes with no predicted function also have signal ratios above 0.5. A further 13 probes returned signal ratios greater than 0.25 at this time point.

Table 5.5. Viral transcripts with signal ratios great than 0.5 in NIH 3T3 cells infected with MHV-68 for 12h.

Probe	Putative Function
M3	Chemokine binding protein (Parry, 2000)
45	Unknown
52	Unknown
M9	Capsid protein
66	Capsid protein
67	Tegument protein
68	Glycoprotein

The change in transcript abundances between 12h and 8h pi is shown in Figure 5.6B. This highlights the changes between the 2 time points, and shows that the majority of transcripts have fallen in abundance. This suggests that few genes were expressed at this stage of the lytic infection cycle. In particular ORF M3, 52 and 59 transcript levels had fallen greatly, although their overall levels were still relatively high. Ten transcripts did increase in abundance between these time points (Table 5.6), including the cluster of predominantly structural genes (ORF M9, 66, 67, 68). The products of these genes are likely to be required by MHV-68 towards the final stages of the replication cycle.

ORF 75C also increased in abundance to a relatively high degree. ORF 75B, which is thought to be another copy of the *N*-formylglycinamide ribotide amidotransferase enzyme encoded by ORF 75C, does not show a similar increase. However, as the 2 ORFs are predicted to share a polyadenylation site, the signal detected by the

ORF 75C probe would be expected to be higher than that by the ORF 75B probe, as is the case for ORFs 58 and 59 (see 5.1.5).

The remaining genes listed in Table 5.6 showed relatively small increases in transcript abundance of +0.05 or less and therefore may not be significant. For example, the probe for M2-M3 is amongst these. The signal ratio for this probe fell by 0.02 between 5h and 8h pi, and then increased by 0.04 between 8h and 12h pi. Overall the M2-M3 probe bound little signal, and its profile showed a gradual increase up to a signal ratio of 0.13 at 12h pi. The generally low signal detected by this probe, and the fluctuation in its expression profile combine to suggest that the signal detected by the M2-M3 probe does not represent a transcript, and most likely represents non-specific hybridisation.

Table 5.6. Transcripts increasing in abundance between 8h and 12h pi by MHV-68.

Probe	Putative Function	Signal Increase
M2-M3	Unknown	0.04
4	Complement regulatory protein (Kapadia, 1999)	0.03
M6	Unknown	0.03
19	DNA packaging protein	0.04
M9	Capsid protein	1.30
66	Unknown	0.86
67	Membrane protein	0.37
68	DNA packaging protein	0.32
M10c	Unknown	0.02
74	G-protein coupled receptor	0.05
75C	*N*-formylglycinamide ribotide amidotransferase	0.19

5.1.7 Array Analysis of Cells Infected for 18h

The transcription profile at 18h pi (Figure 5.7A) was very similar to the profile at 12h pi (Figure 5.6A). The main difference was that the cluster of structural genes (ORF M9, 66, 67, 68) no longer dominated the transcriptome. Instead, their expression levels were now similar to those of ORF M3 and 52. The same 7 genes showed signal ratios greater than 0.5 at 18h pi (Table 5.5) as had done at 12h pi (Table 5.4).

Generally, the trend that was observed between 8h and 12h pi continued to 18h pi; few transcripts increased in abundance. Even the cluster of structural genes that had shown increasing abundance between 8h and 12h pi, was now falling in abundance. This suggests that few structural genes are transcribed after 12h pi, and that by 18h pi, active transcription of viral genes is no longer occurring. However, as overall transcript levels remain high (Figure 5.7A), translation of viral genes is likely to continue.

ORF 52 is highlighted in Figure 5.7B, as it shows a large increase in transcript abundance. However, from 8h pi to 18h pi, its signal ratio decreased by 0.32. Therefore, while there was an increase in signal ratio seen at 18h pi, overall ORF 52 transcript levels were lower at 18h pi than at 8h pi.

5.1.8 Cluster Analysis of Expression Profiles

The analyses performed in 5.1.1 - 5.1.7 are based on the absolute levels of transcript detected by each probe. While quantitative in nature, this approach tends to emphasise highly expressed genes. It is also useful to examine each transcription profile relative to itself, so that features such as the point at which a transcript's abundance peaks, and rate at which that peak is reached, are emphasised. Therefore, the MHV-68 transcription data was converted onto a relative scale so that peak transcript abundance = 1 (2.4.2b). Cluster analysis of the transformed data grouped genes with similar expression profiles (Eisen, 1998). This technique allows groups of genes that are transcribed with similar kinetics but not necessarily at similar levels to be identified. The resulting clusters of genes can be used to make functional predictions for genes that have not been characterised, or that have no characterised homologues (for example ORF 52), as functionally-related genes are often transcribed with similar kinetics (Eisen, 1998).

The results of this analysis were visualised using Treeview software (Eisen, 1998), and are shown in Figure 5.8. Treeview software formats the output as a tree diagram showing the relatedness of expression profiles. To the right of the tree, each gene's expression profile is shown as a series of coloured boxes, with intensity of colour proportional to relative transcript abundance. The gene that the profile represents is given to the far right of Figure 5.8.

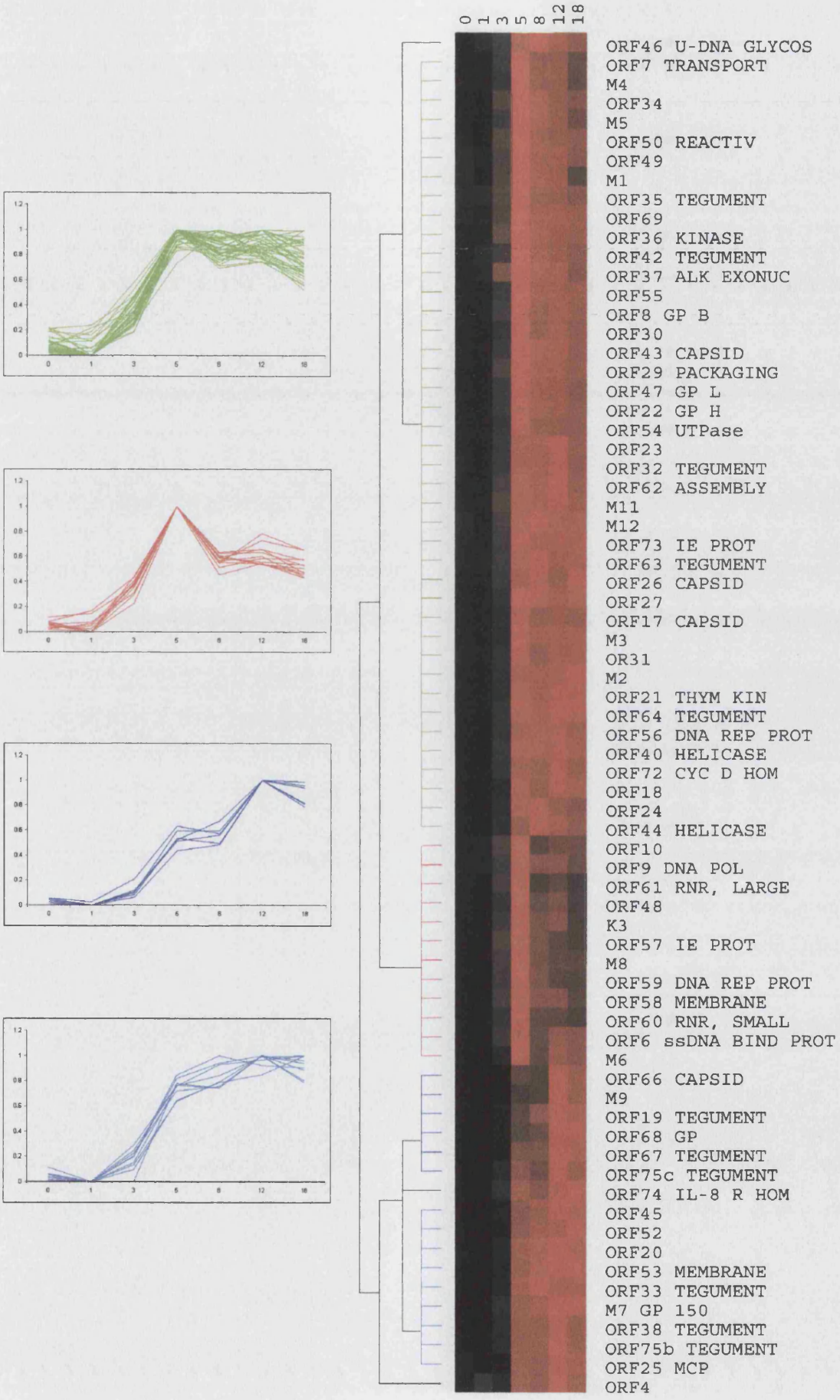
The expression profiles clustered into groups representing 4 general profiles, and these are represented in the 4 line graphs, shown on the left of Figure 5.8. The branches of the tree diagram and the line graphs have been colour coordinated to show which group each graph refers to. It's interesting to note that while protein-

based studies of viral gene expression kinetics generally produce 3 groups (α, β and γ), this RNA-based study has produced 4 groups.

The red cluster consists of transcripts that start to be expressed the earliest, which peak at 5h pi and then fall in abundance. This group of genes includes several genes predicted to be involved in DNA replication (ORFs 6, 9, 59, 60 and 61, r^2 = 0.992), which demonstrates the function-related grouping that results from cluster analysis. The group also contains 3 putative α genes (ORFs K3, M8 and 57, r^2 = 0.997), which suggests that viral DNA replication proteins start to be synthesised at an early stage of the MHV-68 life cycle. This analysis confirms the results of the previous one based on absolute transcript levels (5.1.1 - 5.1.7).

The cluster shown in green showed a peak in expression at 5h pi followed by a more gradual fall, in comparison to the previous cluster (shown in red). This cluster was the largest, containing 41 genes. Functional groupings were also observed within this cluster. For example the genes homologous to glycoprotein L and H of other herpesviruses were placed on adjacent branches of the tree. These 2 glycoproteins are expressed together as the 2 form a heteromeric complex (Hutchinson, 1992). Similarly, the transcription profiles of ORFs 40 and 56 lie adjacent to each other. Both these genes are homologous to characterised genes that are constituents of herpesvirus DNA helicase-primase complexes (Crute, 1989b). ORF 44, which is homologous to the third constituent of the helicase-primase complex, is found to cluster very close to ORFs 40 and 56 as well.

Figure 5.8 (following page)**.** Hierarchical cluster analysis of MHV-68 expression profiles. Data in the form of $\log_2$ normalised means (duplicates, n = 2-6) was converted into a percentage of each gene's maximum. These percentages were imported into Cluster software and hierarchical clustering performed using the average-linkage algorithm and an uncentred correlation similarity matrix. The data is shown as a colour matrix with columns representing time points pi and rows representing each gene's expression profile. Black boxes represent no expression and brighter shades of red correspond to increasing expression. The dendogram shows related expression profiles on the same branch, with branch lengths representing the degree of similarity between individual profiles. Line graphs of clustered genes are shown to the left, with the colour of each line graph corresponding to the dendogram branch of the same hue.


0 1 3 5 8 12 18
ORF46 U-DNA GLYCOS
ORF7 TRANSPORT
M4
ORF34
M5
ORF50 REACTIV
ORF49
M1
ORF35 TEGUMENT
ORF69
ORF36 KINASE
ORF42 TEGUMENT
ORF37 ALK EXONUC
ORF55
ORF8 GP B
ORF30
ORF43 CAPSID
ORF29 PACKAGING
ORF47 GP L
ORF22 GP H
ORF54 UTPase
ORF23
ORF32 TEGUMENT
ORF62 ASSEMBLY
M11
M12
ORF73 IE PROT
ORF63 TEGUMENT
ORF26 CAPSID
ORF27
ORF17 CAPSID
M3
OR31
M2
ORF21 THYM KIN
ORF64 TEGUMENT
ORF56 DNA REP PROT
ORF40 HELICASE
ORF72 CYC D HOM
ORF18
ORF24
ORF44 HELICASE
ORF10
ORF9 DNA POL
ORF61 RNR, LARGE
ORF48
K3
ORF57 IE PROT
M8
ORF59 DNA REP PROT
ORF58 MEMBRANE
ORF60 RNR, SMALL
ORF6 ssDNA BIND PROT
M6
ORF66 CAPSID
M9
ORF19 TEGUMENT
ORF68 GP
ORF67 TEGUMENT
ORF75c TEGUMENT
ORF74 IL-8 R HOM
ORF45
ORF52
ORF20
ORF53 MEMBRANE
ORF33 TEGUMENT
M7 GP 150
ORF38 TEGUMENT
ORF75b TEGUMENT
ORF25 MCP
ORF4

The remaining 2 clusters are shown in light and dark blue in Figure 5.8. The genes in these 2 clusters (those with predicted functions) are homologous to herpesvirus structural genes. The light blue cluster contains ORFs 25 (major capsid protein), 33 (tegument protein), 38 (membrane protein), M7 (glycoprotein 150; Stewart, 1996) and 53 (membrane protein). It also contains ORFs 20, 45 and 52, which have no characterised homologues. The dark blue cluster includes ORFs 19 (tegument), 66 (capsid) and 68 (glycoprotein). ORF M9 is one of the genes in this cluster and was originally designated as unique to MHV-68 (Virgin, 1997). However, it has since been found to be homologous to a capsid protein (ORF 65) of other gammaherpesviruses (Milligan, 1998).

The key difference between the expression patterns of these 2 clusters is the time at which they reach their peak abundance. The light blue cluster reaches and maintains a peak of expression at around 8h pi. The dark blue cluster tends to peak around 12h pi and are the last set of genes to reach peak levels of expression. It would be interesting to characterise the genes further to see if the 2 clusters correlate to the γ1 and γ2 kinetic classes.

These results highlight the ability of cluster analysis to produce functionally-related groups of genes. However, there are some drawbacks to this technique. In particular the potential of cluster analysis is somewhat limited by the nature of herpesvirus gene expression, as herpesviruses characteristically transcribe their genes in a tightly regulated cascade. Therefore, genes of the same kinetic class are likely to have similar expression profiles. This can be seen in the green cluster, which includes many functional classes of genes. They include the putative genes for thymidine kinase (ORF 21), glycoprotein H (ORF 22) and a cyclin D homologue (ORF 72). These functionally diverse genes cluster together, although within the cluster there are more related groups of genes. Perhaps, a more rigorous algorithm to group expression profiles could be useful to further divide this group. Viral genes tend to all be expressed with similar kinetics and so an algorithm specifically designed for analysis of viral transcription could prove useful in this case. The development of further analytical techniques for array data remains a priority for genomic studies.

Cluster analysis of MHV-68 gene expression has shown that different functional groups can be temporally related. Therefore, while it is possible to hypothesise that genes without characterised homologues in the red cluster (ORFs 10, 48 and M6) could be involved in DNA replication or α genes, like the other cluster members, further study is required before more robust predictions can be made. M6 is

particularly interesting as it is a gene unique to MHV-68, with no homologues currently in the gene databases.

5.2 Gene Expression in the Presence of Viral Protein Synthesis Inhibition

5.2.1 Inhibition of Protein Synthesis with Cycloheximide

To dissect the transcriptional behaviour of MHV-68 further, α genes were characterised by blocking *de novo* protein synthesis with CX. Initial studies using 100μg/ml and 50μg/ml concentrations of CX resulted in cell death of NIH 3T3 cell monolayers, although these concentrations were often used in similar studies (Nicholas, 1990; Haque, 2000). However, other studies have shown that these levels of CX can result in apoptosis for a variety of cell types (Martin, 1995). Therefore, to define the concentration required for 95% inhibition of protein synthesis, the metabolism of radiolabelled methionine by NIH 3T3 cells was measured in the presence of 100μg/ml – 0.1ng/ml CX.

NIH 3T3 cell monolayers were incubated in the presence of radiolabelled methionine and a range of CX concentrations. After 6h, the incorporation of methionine was quantified. Figure 5.9 shows that in the presence of 50μg/ml and 100μg/ml, the NIH 3T3 cells were unable to metabolise any methionine. In fact, 95% inhibition of protein synthesis was obtained with only 2μg/ml CX, and therefore this dose was used in all subsequent experiments. As the CX was solublised in DMSO, control monolayers were incubated with equivalent concentrations of DMSO alone. These cells showed no inhibition of methionine uptake.

5.2.2 Characterisation of MHV-68 α Genes

An array experiment was set up to investigate gene expression in the presence of 95% protein synthesis inhibition. NIH 3T3 cell monolayers were both pre-treated with, and then infected in the presence of CX. Negative controls (uninfected cells in the presence of CX) and positive controls (cells infected in the absence of CX) were also included. Cells were harvested at 5h and 8h pi and RNA was isolated for subsequent array hybridisations. The results of these hybridisations are shown in Figure 5.10.

At both 5h and 8h pi, ORF 73 transcript levels were higher when protein synthesis was inhibited, than when it was not. This suggested that ORF 73 continued to be transcribed, and accumulated, in the absence of a switch to β-class gene

expression, and that ORF 73 transcription did not require *de novo* viral protein synthesis. Therefore, ORF 73 was characterised as an α gene.

At 5h pi, ORFs 29b and 35 showed slightly higher transcript abundances when β gene transcription was inhibited by CX than when it was not. However, at 8h pi, the levels of these genes had decreased compared to their levels at 5h pi when *de novo* protein synthesis was inhibited. Furthermore their transcript levels increased between 5h and 8h pi if CX was not used. This suggests that these genes are not α genes as the opposite trend in their levels would be expected if this was the case. Indeed, examining the results for uninfected samples hybridised to the array show that the signals for these 2 ORFs were similar to background levels, which again suggests that these are not α genes (data not shown).

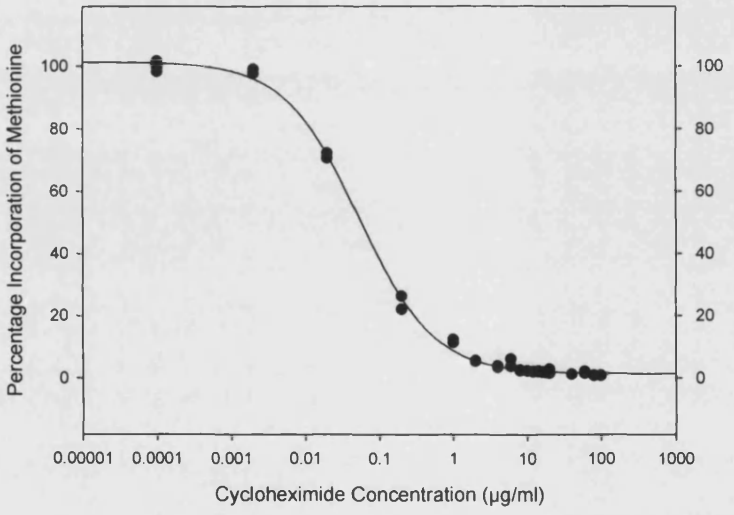


Figure 5.9. Incorporation of radiolabelled methionine by NIH 3T3 cells in the presence of CX.

NIH 3T3 cell monolayers were incubated in the presence of varying concentrations of CX. The cells were then infected with MHV-68 and radiolabelled methionine introduced into the medium. Cells were harvested after 6h and the level of methionine incorporation measured. The y-axis shows percentage incorporation of labelled methionine relative to uninhibited controls. The x-axis shows the concentration of CX in the cell medium.

Fig 5.10A. Cells infected for 5h in the presence of CX

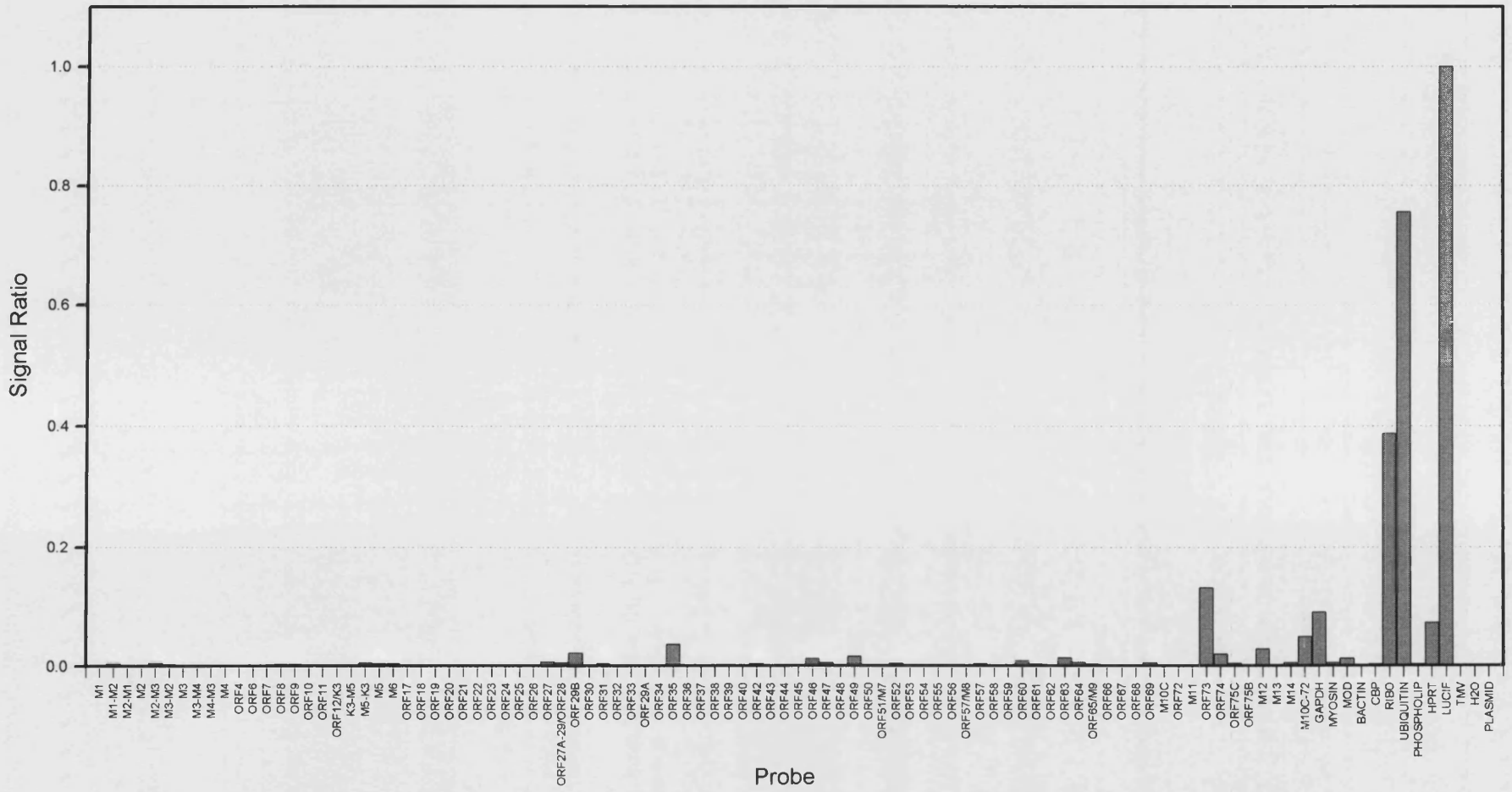


Fig 5.10B. Cells infected for 5h without inhibition

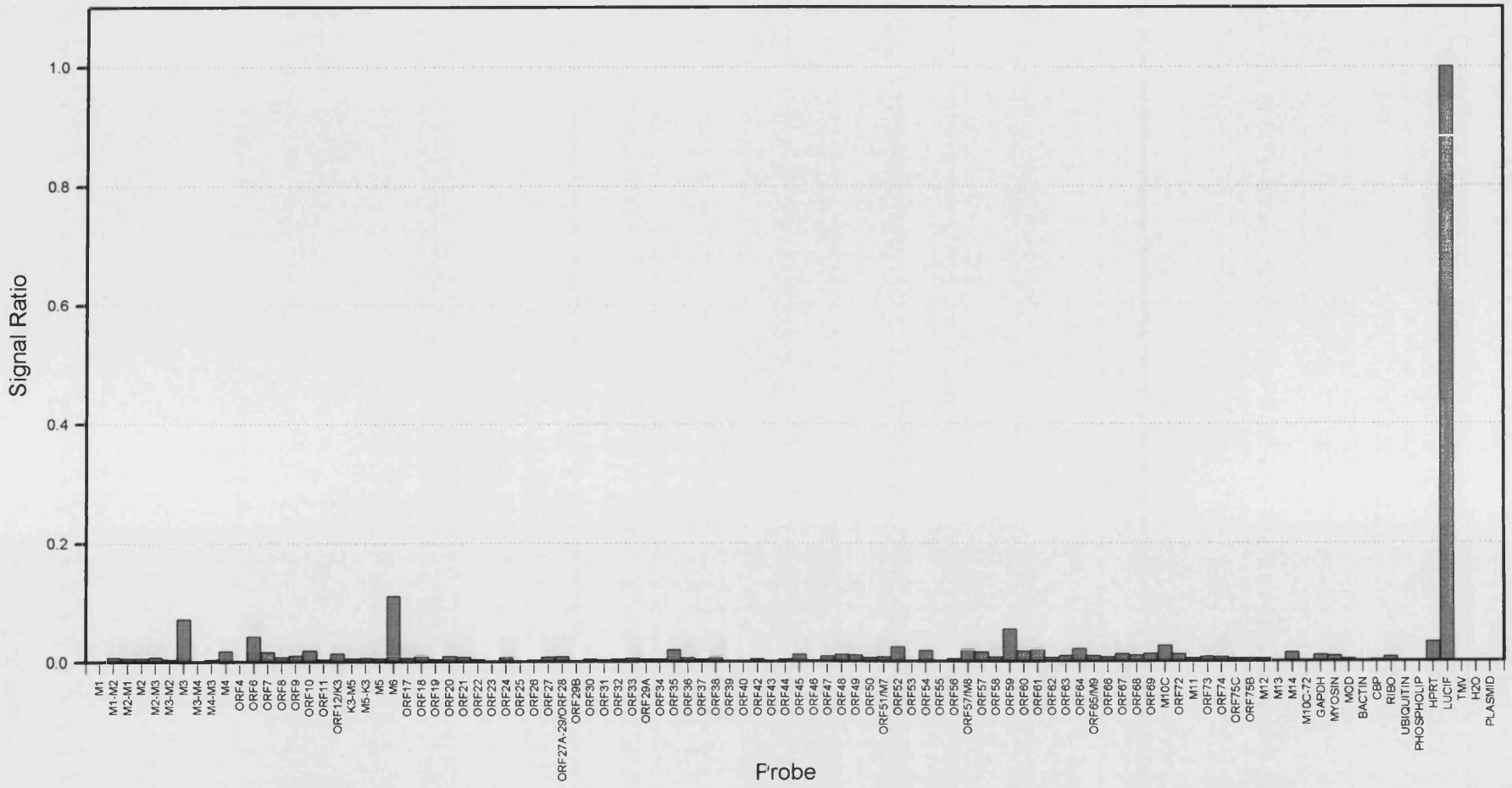


Fig 5.10C. Cells infected for 8h in the presence of CX

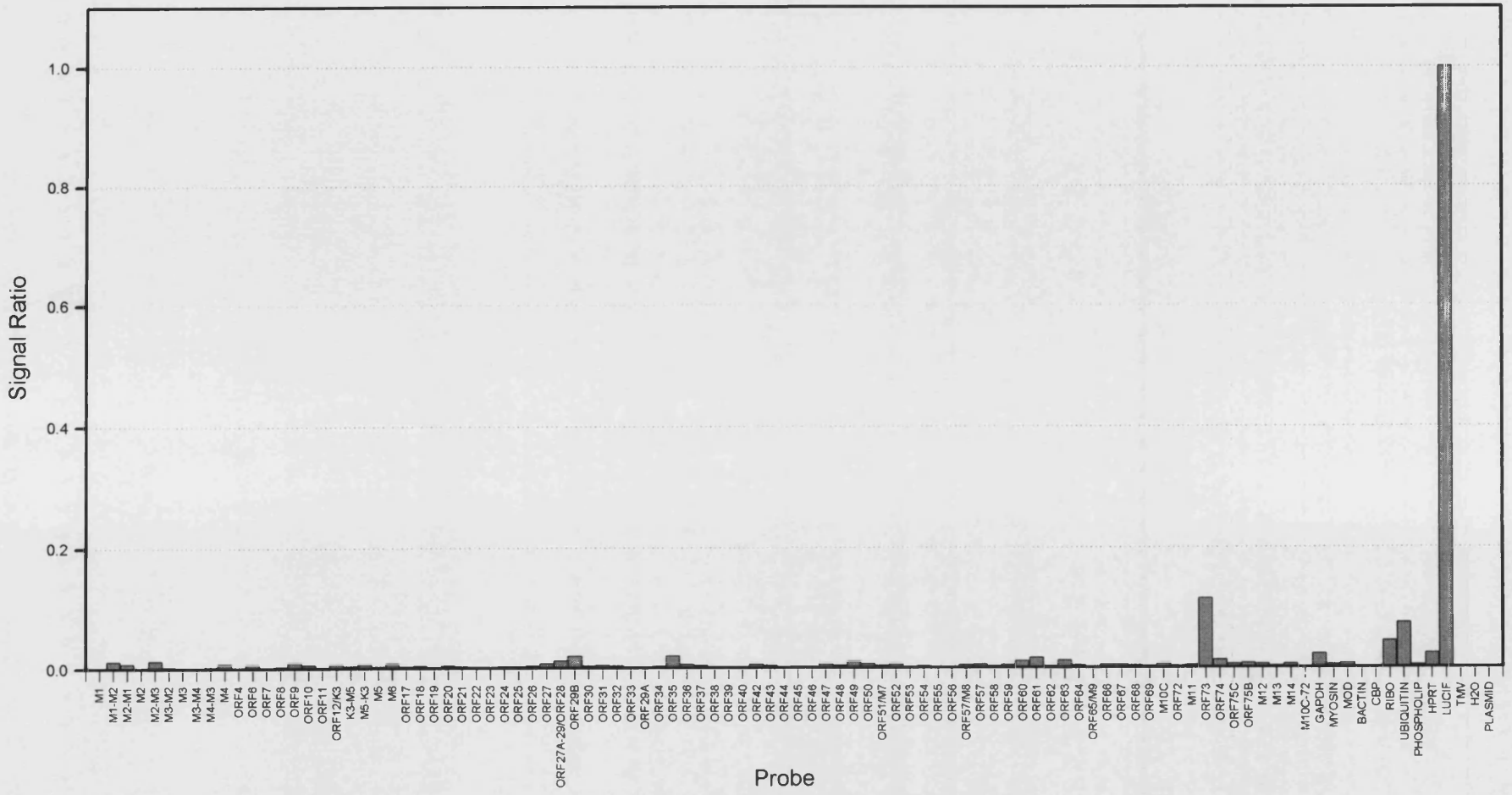


Fig 5.10D. Cells infected for 8h without inhibition

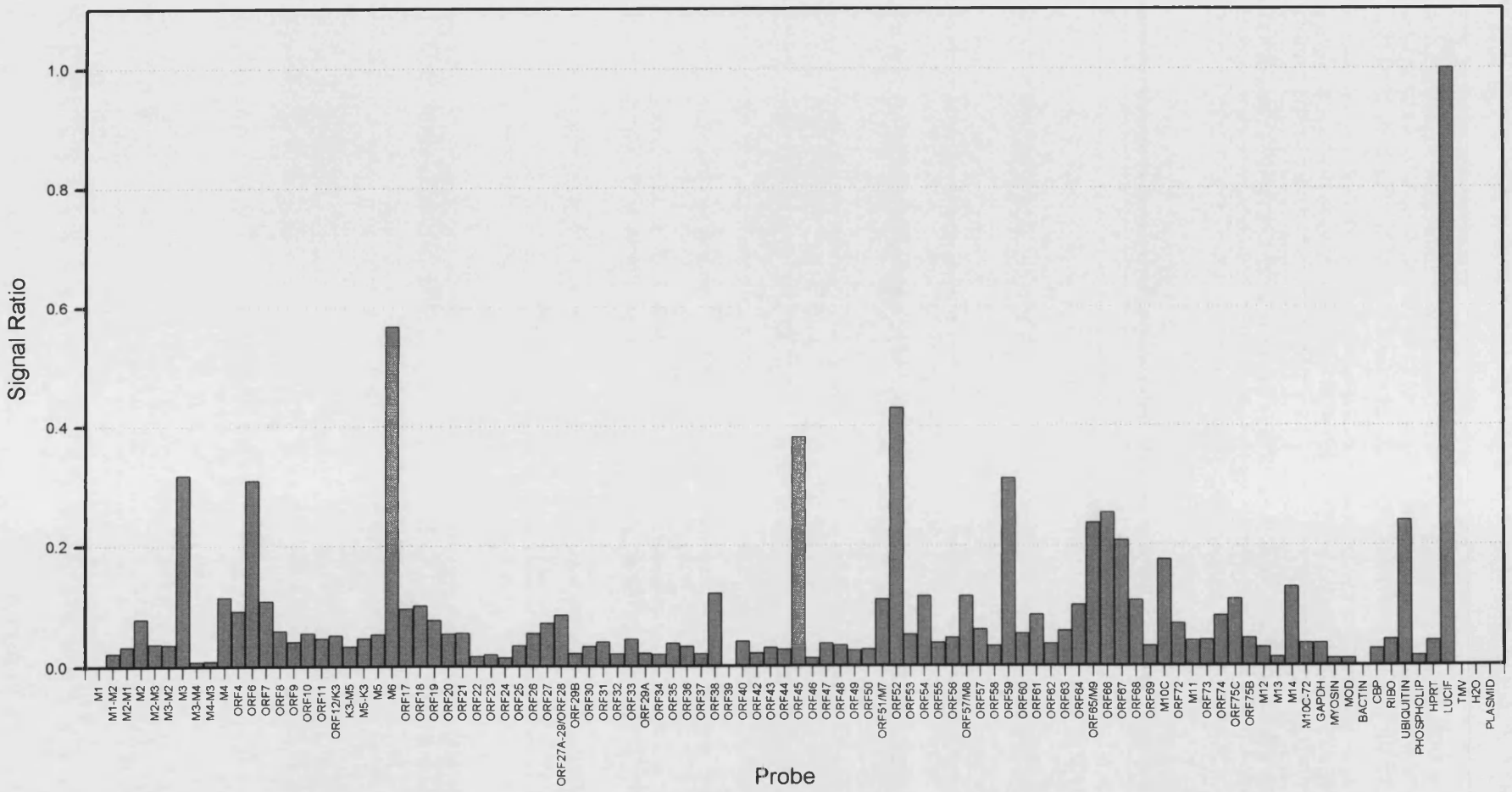


Figure 5.10 (preceding pages). DNA array analysis of MHV-68 gene expression in the absence of *de novo* protein synthesis.
NIH 3T3 cells were infected in the presence of CX to produce a 95% protein synthesis block, or without inhibition. RNA was isolated from cells harvested at 5h and 8h pi, and used to produce radiolabelled target for hybridisation to arrays. Signals from a single representative experiment were quantified, normalised and plotted as bar charts. **A.** Array analysis of cells infected in the presence of CX 5h pi. **B.** Array analysis of uninhibited cells 5h pi. **C.** Array analysis of cells infected in the presence of CX 8h pi. **D.** Array analysis of uninhibited cells 8h pi.

5.3 Gene Expression in the Presence of Viral DNA Replication Inhibition

Having characterised α gene expression, γ genes were also studied. 4'-S-EtdU is a thionucleoside analogue and has been shown to be a potent inhibitor of MHV-68 DNA replication (Barnes, 1999). Therefore, NIH 3T3 cells were pre-treated and infected in the presence of 200ng/ml 4'-S-EtdU, which has been shown to completely inhibit MHV-68 replication as determined by plaque assay (Barnes, 1999). Negative (uninfected with 4'-S-EtdU) and positive (infected without inhibition) controls were also set up. Cells were harvested at 5h and 18h pi.

The results of hybridisations with samples taken at 18h pi (n = 4) are shown in Figure 5.11. The transcription profile of an infection where viral DNA replication has been inhibited by the antiviral 4'-S-EtdU is shown in Figure 5.11A, and that of an uninhibited infection in Figure 5.11B. It is clear that there is reduced viral transcription globally in the presence of 4'-S'EtdU.

Figure 5.11C quantifies the difference between these 2 profiles, as the uninhibited dataset has been subtracted from the inhibited dataset and the results plotted. Negative values (represented as bars below the x-axis) indicate reduced transcript abundances when viral DNA replication is blocked, while positive values (represented by bars rising above the x-axis) represent transcripts that accumulate when the block is in place. This shows that the presence of 4'-S-EtdU has effectively reduced the transcription of some transcripts, e.g. ORF 52 and the cluster of structural genes (ORF M9 – 68) that are abundant at this time point. Its presence has also resulted in some transcripts increasing in abundance relative to uninhibited infections, indicating that blocking viral DNA replication has resulted in continued transcription and accumulation of some transcripts.

Fig 5.11A. Cells infected for 18h in the presence of 4'-S-EtdU

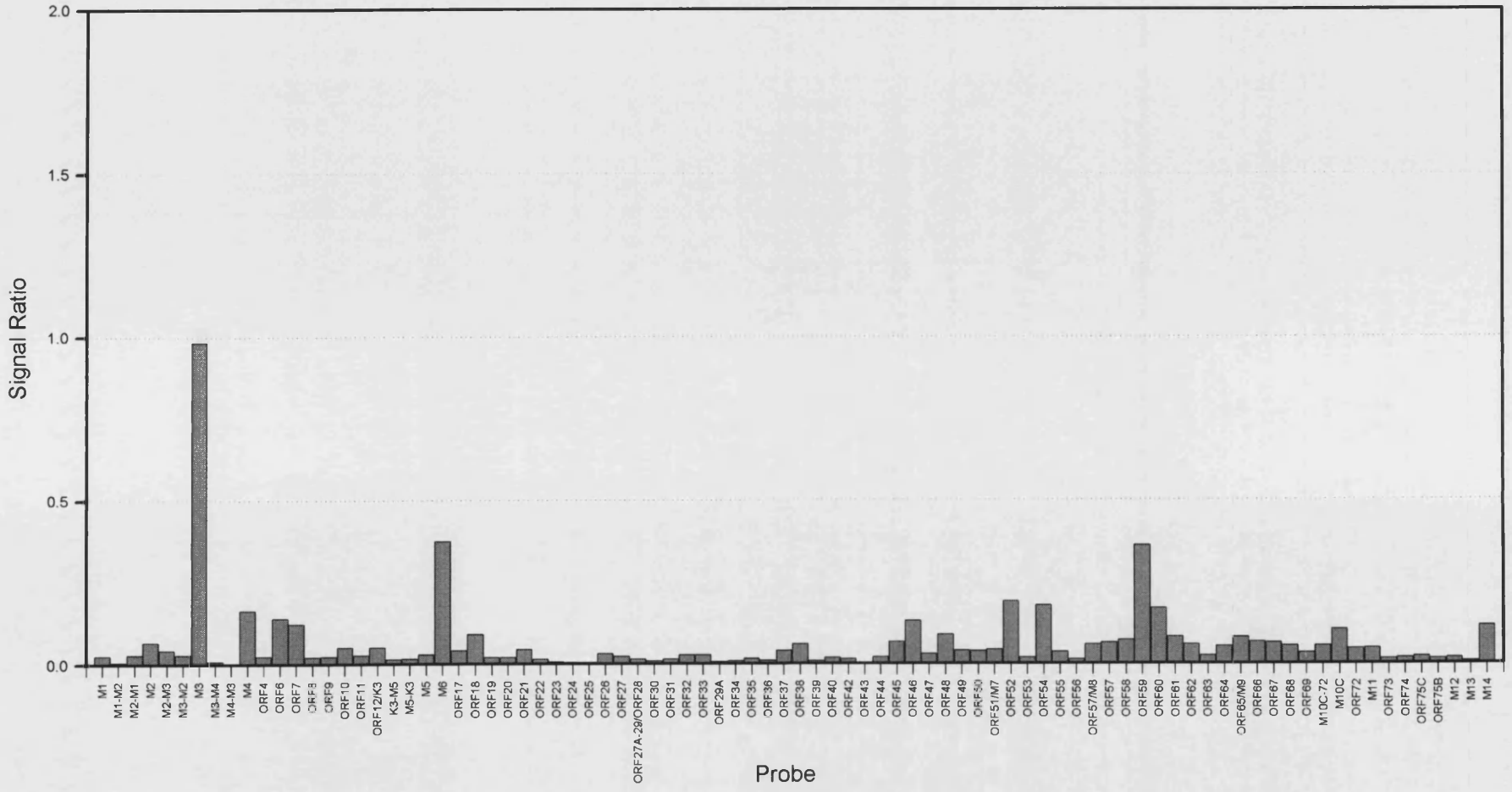


Fig 5.11B. Cells infected for 18h without inhibition

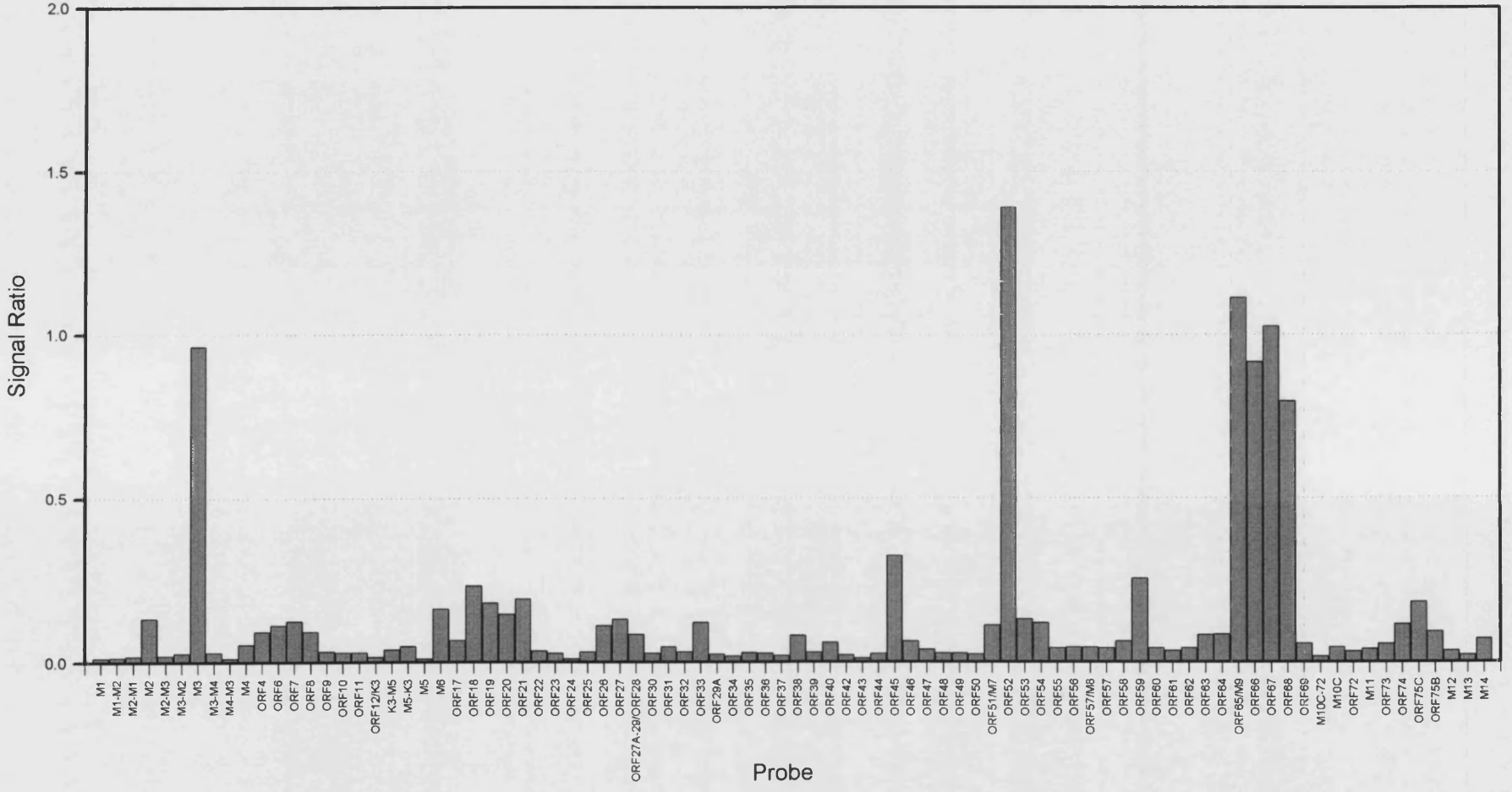


Fig 5.11C. Difference in signal ratios between inhibited and uninhibited infections

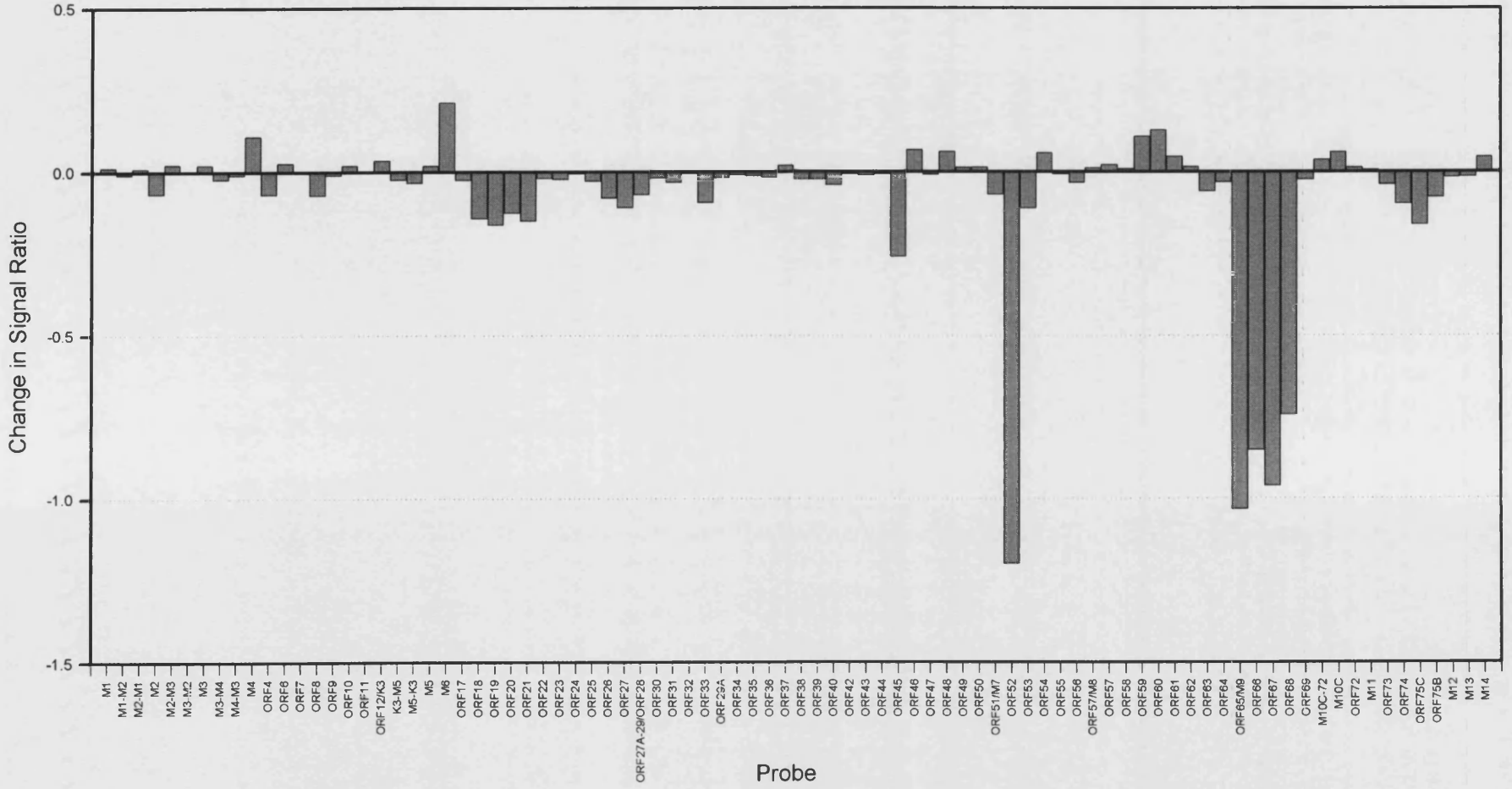


Fig 5.11D. The relative difference in transcript abundances between inhibited and uninhibited infections

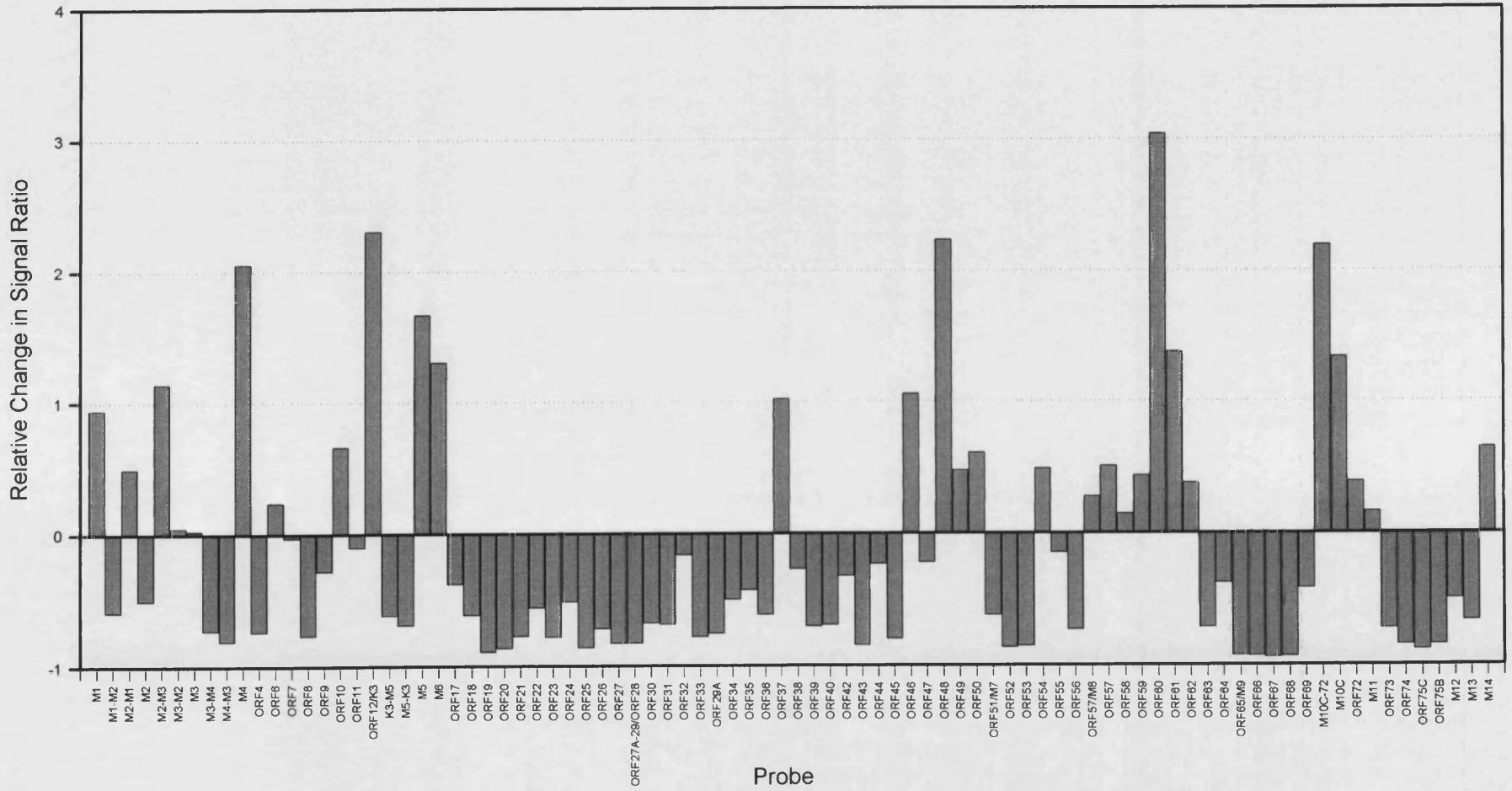


Figure 5.11 (preceding pages). Expression of MHV-68 genes in the absence of viral DNA replication.
NIH 3T3 cells were infected with or without inhibition of viral DNA replication. RNA was isolated from cells harvested 18h pi and used to produce radiolabelled target that were hybridised to arrays. Array data was quantified using a phosphoimager and the results plotted as bar charts (n = 4). A. Array analysis of cell infected for 18h in the presence of 4'-S-EdtU. B. Array analysis of uninhibited cells 18h pi. C. Difference in signal ratios due to inhibition of viral DNA replication. The dataset from the uninhibited infection was subtracted from the inhibited infection dataset. The results were plotted on a bar chart so that an increase in transcript abundance following inhibition is represented by bars above the x-axis. Bars below the x-axis represent decreased transcript abundance following inhibition. D. The relative difference in signal ratios due to inhibition of viral DNA replication. The differences between the uninhibited and inhibited datasets was transformed onto a relative scale so that the change in signal for each probe was represented as a percentage of its signal when infection was uninhibited. Y-axis values of +1 correspond to a 100% increase in transcript abundance, and –1 to a 100% reduction in transcript abundance.

As different transcripts are present at different levels, the results of Figure 5.11C were transformed onto a relative scale to further analyse the impact of the viral DNA replication block on each transcript's levels relative to itself. These results are shown in Figure 5.11D. Again bars below the x-axis show reduced levels and bars above the axis, increased levels of transcripts. However, a value of –1 on the y-axis indicates that the transcript's abundance was reduced to zero by the presence of 4'-S-EtdU. A value of +1 on the y-axis indicates that the abundance of the transcript has doubled in the presence of 4'-S-EtdU. The absence of a bar, or a value of 0 on the y-axis, indicates that there was no difference in expression between uninhibited and inhibited infections.

Complete inhibition of viral DNA replication has been observed with 200ng/ml 4'-S-EtdU (Barnes, 1999), but Figure 5.11D shows a less than complete shutoff of many potential γ genes occurring under the same conditions. While most probes showed close to 100% reduction in signal when viral DNA replication was inhibited, several probes showed a range of reductions in transcript abundance. Forty-two probes showed a reduction in transcript abundance of 50% or more, and 15 probes showed less than 50% signal reduction. The remaining 29 probes showed an increase in signal, and were therefore designated β genes.

To ascertain which transcripts were significantly reduced in abundance, in the presence of 4'-S-EtdU, the standard deviation for probe signals was calculated from

4 replicate arrays. Only signals that became reduced by more than 2 measures of standard deviation were characterised as those of γ genes. Forty-nine probes met these criteria, as listed in Table 5.7, and included predicted ORFs as well as inter-genic regions.

The inclusion of inter-genic probes in this group is interesting, as this suggested that these probes are detecting transcripts whose abundance are affected by the inhibition of viral transcription. However, as the analysis shown in Figure 5.11D is a relative one, it is useful to refer back to Figure 5.11B and Figure 5.11C to see the actual levels of signal detected by these probes and also the quantitative effects of 4'-S-EtdU inhibition.

Figure 5.11B shows that the level of signal detected by the inter-genic probes in an uninhibited infection are low (signal ratios ranging from 0.014 – 0.048). Figure 5.11C shows that the reduction in signal when viral DNA replication is inhibited is also low (reductions range from 0.008 – 0.033). These results indicate that the change in signal for these inter-genic probes is high relative to themselves, but very small quantitatively. Therefore, further studies are required to determine whether these inter-genic regions do encode transcripts.

One inter-genic probe did however show clear relative and absolute changes in signal when viral DNA replication was inhibited. This was the probe for ORF 27-29A, which showed an 82% reduction in signal, or a change of –0.07 in absolute terms. This suggested that this inter-genic probe could be detecting a transcript.

Further analysis showed that a number of genic probes also showed very small absolute changes, but which corresponded to large relative changes. Therefore, ORFs 21, 36, 40, 56 and 73 were designated non-γ genes and are shown in a grey font in Table 5.7. The remaining 38 genes were designated γ genes and are shown in a black font.

Interestingly, even at 5h pi, there was a difference in signal detected for some of the γ genes between inhibited and uninhibited infections (data not shown). Similar to the inter-genic probes at 18h pi, while the relative differences were high, the quantitative analyses show that the actual signals were small and the difference in signals smaller. Therefore, further study is required before this difference can be confirmed.

Table 5.7. Probes showing reduced signal following inhibition of viral DNA replication.

Probe	Putative Function	Change in Transcript Levels
M1-M2	Unknown	-59%
M2	Unknown	-50%
M3-M4	Unknown	-73%
M4-M3	Unknown	-81%
4	Complement regulatory protein	-74%
8	Glycoprotein B	-77%
K3-M5	Unknown	-62%
M5-K3	Unknown	-69%
17	Capsid protein	-38%
18	Unknown	-61%
19	Minor capsid protein	-89%
20	Unknown	-86%
21	Thymidine kinase	-77%
22	Glycoprotein H	-56%
23	Unknown	-78%
24	Unknown	-51%
25	Major capsid protein	-86%
26	Capsid protein	-72%
27	Unknown	-83%
27A-29	Unknown	-82%
29	DNA packaging protein	-76%
30	Unknown	-68%
31	Unknown	-69%
33	Tegument protein	-77%
34	Unknown	-50%
35	Tegument protein	-43%
36	Serine-threonine protein kinase	-61%
39	Glycoprotein M	-70%
40	Component of helicase-primase complex	-69%
42	Tegument protein	-32%
43	Minor capsid protein	-84%
45	Unknown	-80%
M7	Glycoprotein 150	-62%
52	Unknown	-86%
53	Membrane protein	-85%
56	Component of helicase-primase complex	-73%
63	Tegument protein	-71%

Probe	Putative Function	Change in Transcript Levels
64	Tegument protein	-38%
M9	Capsid protein	-93%
66	Capsid protein	-93%
67	Tegument protein	-94%
68	DNA packaging protein	-93%
69	None	-42%
73	None	-72%
74	G-protein coupled receptor	-85%
75c	Formylglycinamide ribotide amidotransferase	-89%
75b	Formylglycinamide ribotide amidotransferase	-85%
M12	None	-50%
M13	None	-66%

5.4 Verification of Array Data

Some of the array data presented here is reflected and confirmed in previous work using alternate methods, such as northern blot analysis and RNAse protection assay, to examine MHV-68 transcription (Mackett, 1997; Husain, 1999; Simas, 1999; Wu, 2000b; Rochford, 2001). However, northern blot analysis was also employed here to further confirm the data presented. Northern blots were prepared with RNA isolated from MHV-68 infected NIH 3T3 cells at various times pi, along with mock-infected controls. Additional northern blots were prepared using RNA isolated from cells infected in the presence of CX to represent α gene expression, and also 4'-S-EtdU to show γ gene expression.

Ten μg of total RNA, which had been spiked with 10ng luciferase mRNA, was loaded into each well of a denaturing agarose gel. Once RNA samples had been size separated by electrophoresis, they were blotted onto membranes. All probes were sequence verified before hybridisation to blots.

5.4.1 Northern Blot Analysis of ORF M3

When northern blots were probed for ORF M3 (chemokine binding protein), a single transcript of approximately 1.2kb was detected. M3 was expressed at 5h pi and showed continued high levels of mRNA up to 18h pi, as shown in Figure 5.12A. This confirmed the results obtained from array analyses (Figure 5.2 - Figure 5.7).

ORF M3 transcripts could not be detected when cells were infected in the presence of CX, as shown in Figure 5.13A. This confirmed that M3 was not an α gene as concluded from the array analyses shown in Figure 5.10.

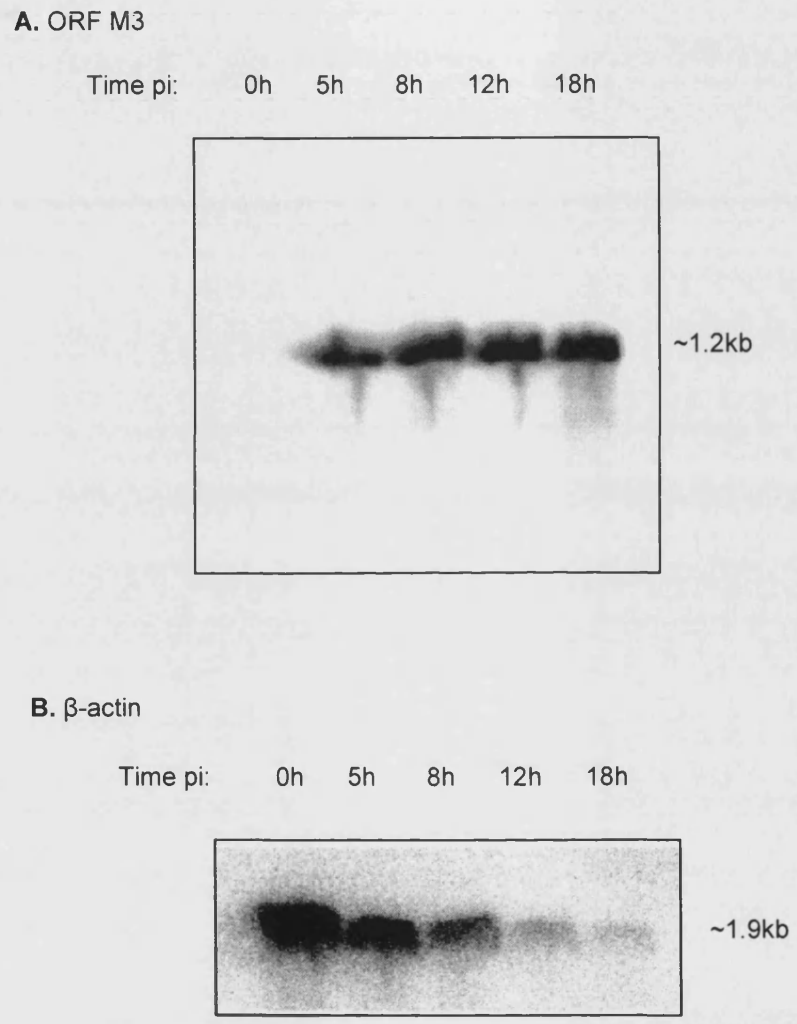
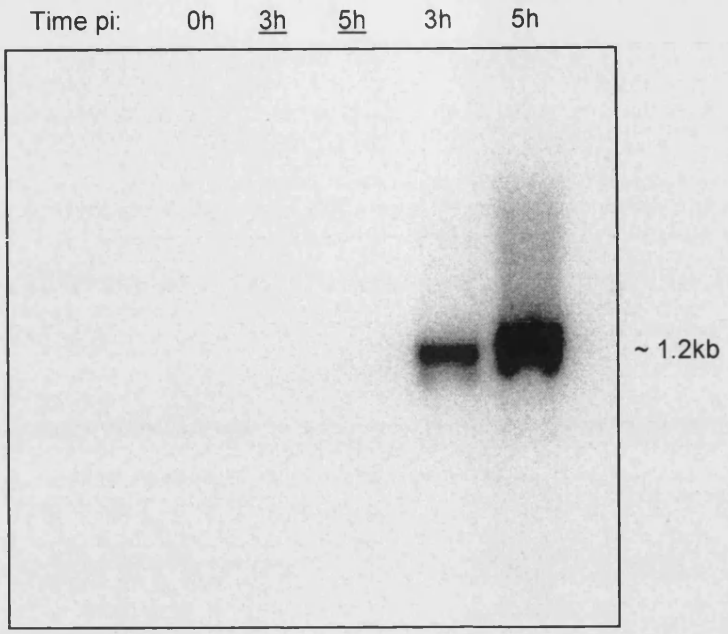


Figure 5.12. Northern blot analysis of ORF M3 transcription.
NIH 3T3 cells were infected and harvested for isolation of RNA at various time points pi. The RNA was size separated on a denaturing gel and blotted onto nylon membranes. Radiolabelled probes were then hybridised to the membranes. The results were visualised using a phosphoimager. A. Hybridisation with ORF M3 specific probe. B. A control hybridisation for β-actin. The time pi when cells were taken for RNA isolation is indicated above each panel. Approximate sizes of transcripts are indicated to the right.

A. ORF M3



B. β-actin

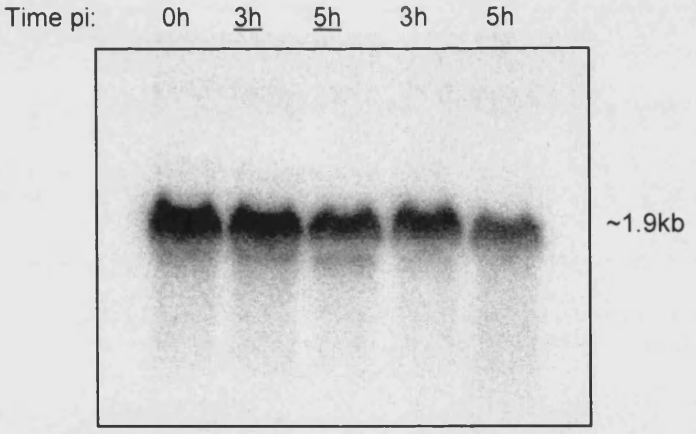


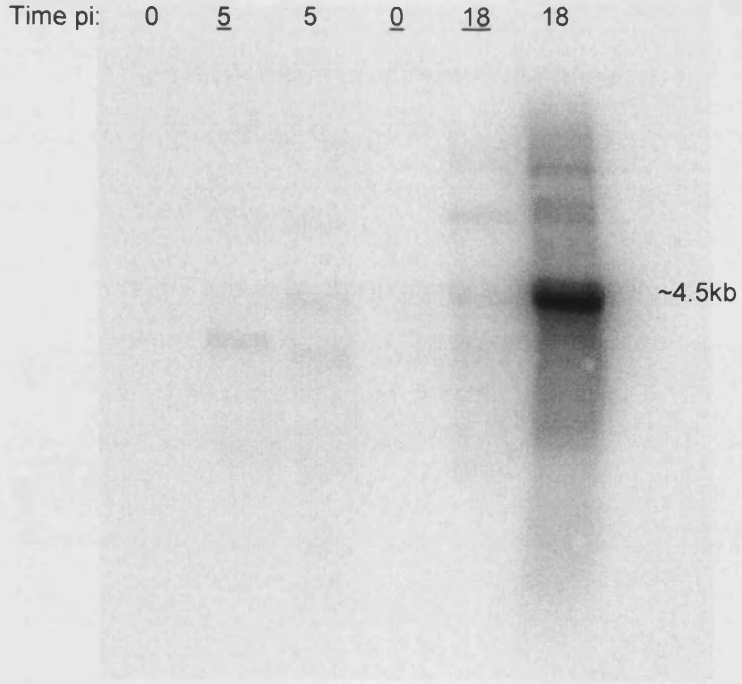
Figure 5.13. Northern blot analysis of ORF M3 transcripts in the presence of CX.
NIH 3T3 cells were infected and harvested for isolation of RNA at various time points pi. The RNA was size separated on a denaturing gel and blotted onto nylon membranes. Radiolabelled probes were then hybridised to the membranes. The results were visualised using a phosphoimager. **A.** Hybridisation with ORF M3 specific probe. **B.** Control hybridisation for β-actin. The time pi when cells were taken for RNA isolation is indicated above each panel. Underlined time points indicate presence of CX. Approximate sizes of transcripts are indicated to the right.

5.4.2 Northern Blot Analysis of ORF 67

A probe for ORF 67 (tegument protein) was found to hybridise to RNA isolated 18h pi, predominantly to a band of around 4.5kb. This hybridisation is shown in Figure 5.14A. There were also a number of faint bands at 18h pi. This was expected as ORF 67 has been predicted to be part of a nested set of co-terminal genes with ORF 65 and 66 (Milligan, 1998). Very faint bands were also visible at 5h pi. Transcription of ORF 67, 66 and 65 were detected by the array from 5h to 18h pi, as shown in Figure 5.4 - Figure 5.7.

ORF 67 probe was also hybridised to RNA isolated from cells infected in the presence of 4'-S-EtdU, also shown in Figure 5.14A. At 18h pi, the predominant band was found to be greatly reduced in intensity, suggesting that the transcript showed γ kinetics. Fainter bands also appeared reduced in intensity. Characterisation of γ genes with the array had concluded that ORF 67, 66 and 65 were all γ genes, as shown in Figure 5.11.

A. ORF 67



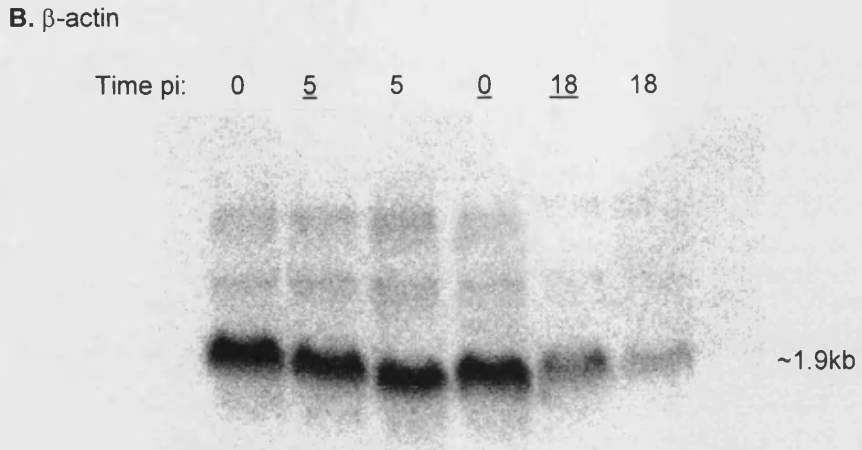


Figure 5.14 (preceding and current page). Northern blot analysis of ORF 67 transcripts in the presence of 4'-S-EtdU.
NIH 3T3 cells were infected and harvested for isolation of RNA at various time points pi. The RNA was size separated on a denaturing gel and blotted onto nylon membranes. Radiolabelled probes were then hybridised to the membranes. The results were visualised using a phosphoimager. A. ORF 67 specific probe B. β-actin. The time pi when cells were taken for RNA isolation is indicated above each panel. Underlined time points indicate presence of 4'-S-EtdU. Approximate sizes of transcripts are indicated to the right.

5.4.3 Northern Blot Analysis of ORF 52 and 53

ORF 52 and 53 are also predicted to share a poly-adenylation signal (Milligan, 1998). An ORF 52 probe was found to hybridise predominantly to a band of around 400bp, and also to a band of around 700bp, as shown in Figure 5.15A. The 400bp band was visible at 5h pi, but was much stronger between 8h and 18h pi. The 700bp band was fainter than the smaller one, and present between 8h and 18h pi. The array showed that ORF 52 was expressed from 5h pi, and was heavily expressed between 5h and 18h pi (Figure 5.4 - Figure 5.7).

A probe for ORF 53 was also hybridised to the same blot. This hybridised to a single band of around 700bp, as shown in Figure 5.15B. A faint band was seen at 8h, 12h and 18h pi. Array analysis of ORF 53 transcription suggested that ORF 53 was expressed at lower levels than ORF 52. The array showed that ORF 53 started to be expressed at 5h pi and was then expressed at higher levels up to 18h pi (Figure 5.4 - Figure 5.7).

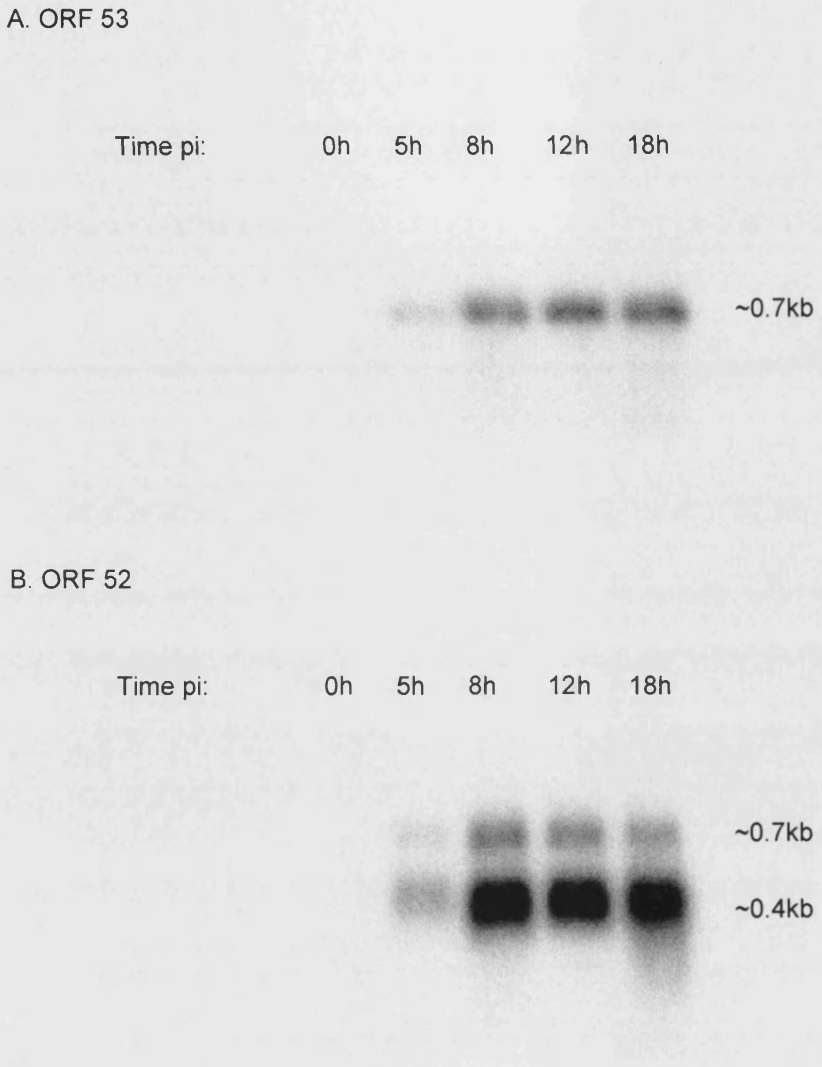


Figure 5.15. Northern blot analysis of ORF 52 and 53 transcripts.
NIH 3T3 cells were infected and harvested for isolation of RNA at various time points pi. The RNA was run out on a denaturing gel and blotted onto nylon membranes. Radiolabelled probes were then hybridised to the membranes. The results were visualised using a phosphoimager. A. ORF 53 probe. B. ORF 52 probe. It should be noted that the same blots were stripped and reprobed in Figure 5.12 and Figure 5.15. The time pi when cells were taken for RNA isolation is indicated above each panel. Approximate sizes of transcripts are indicated to the right.

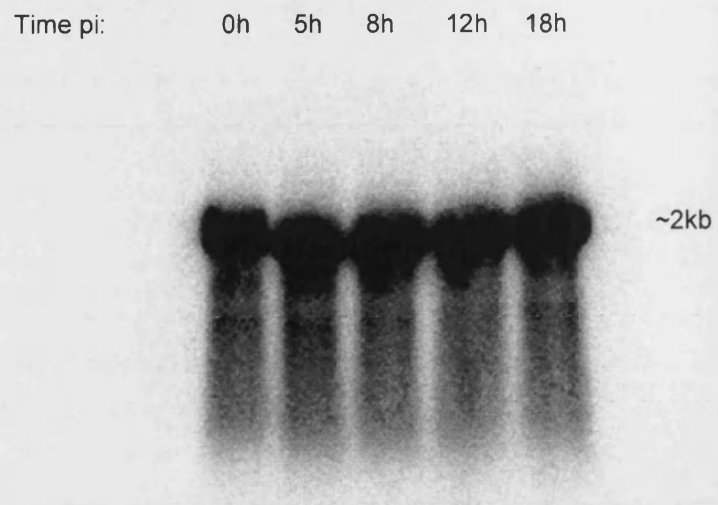
5.4.4 Northern Blot Analysis of β-actin

Each of the northern blots probed for viral transcripts, were stripped and probed for β-actin. These control hybridisations showed that the abundance of β-actin transcripts reduced over the course of infection with MHV-68. By 18h pi, β-actin transcripts were hardly detectable, as shown in Figure 5.12B. This result confirmed the observation made from initial array analyses that showed housekeeping gene transcripts were down-regulated following infection. This also reconfirmed the need for a control, additional to the housekeeping genes, for normalisation of array datasets allowing their cross-comparison.

Luciferase mRNA was also probed for, as shown in Figure 5.16. This showed a band of equal intensity in all lanes, which suggested that equal quantities of luciferase mRNA was present in each lane. Therefore, equal quantities of luciferase mRNA had been spiked into each RNA sample.

β-actin transcript levels were reduced in cells infected in the presence of CX, relative to levels in mock-infected cells, as shown in Figure 5.13B. However, the levels of β-actin appeared less affected when the infection was inhibited by CX, relative to uninhibited infections. This suggested that while host protein synthesis shut off occurred in the absence of viral protein synthesis, it was most efficient when MHV-68 α proteins were produced. Finally, the relative abundance of β-actin in the uninfected lanes suggested that any toxicity to the cells caused by CX was minimal.

Figure 5.16 (on following page). Northern blot analysis to confirm internal luciferase control. NIH 3T3 cells were infected and harvested for isolation of RNA at various time points pi. Ten ng luciferase RNA spiked into each sample. RNA was then size separated on a denaturing gel and blotted onto nylon membranes. Radiolabelled probes were then hybridised to the membranes. The results were visualised using a phosphoimager. It should be noted that this blot is the same blot used in Figure 5.12 and Figure 5.16 (after stripping). The time pi when cells were taken for RNA isolation is indicated above each panel.



5.5 Summary of MHV-68 Lytic Gene Expression

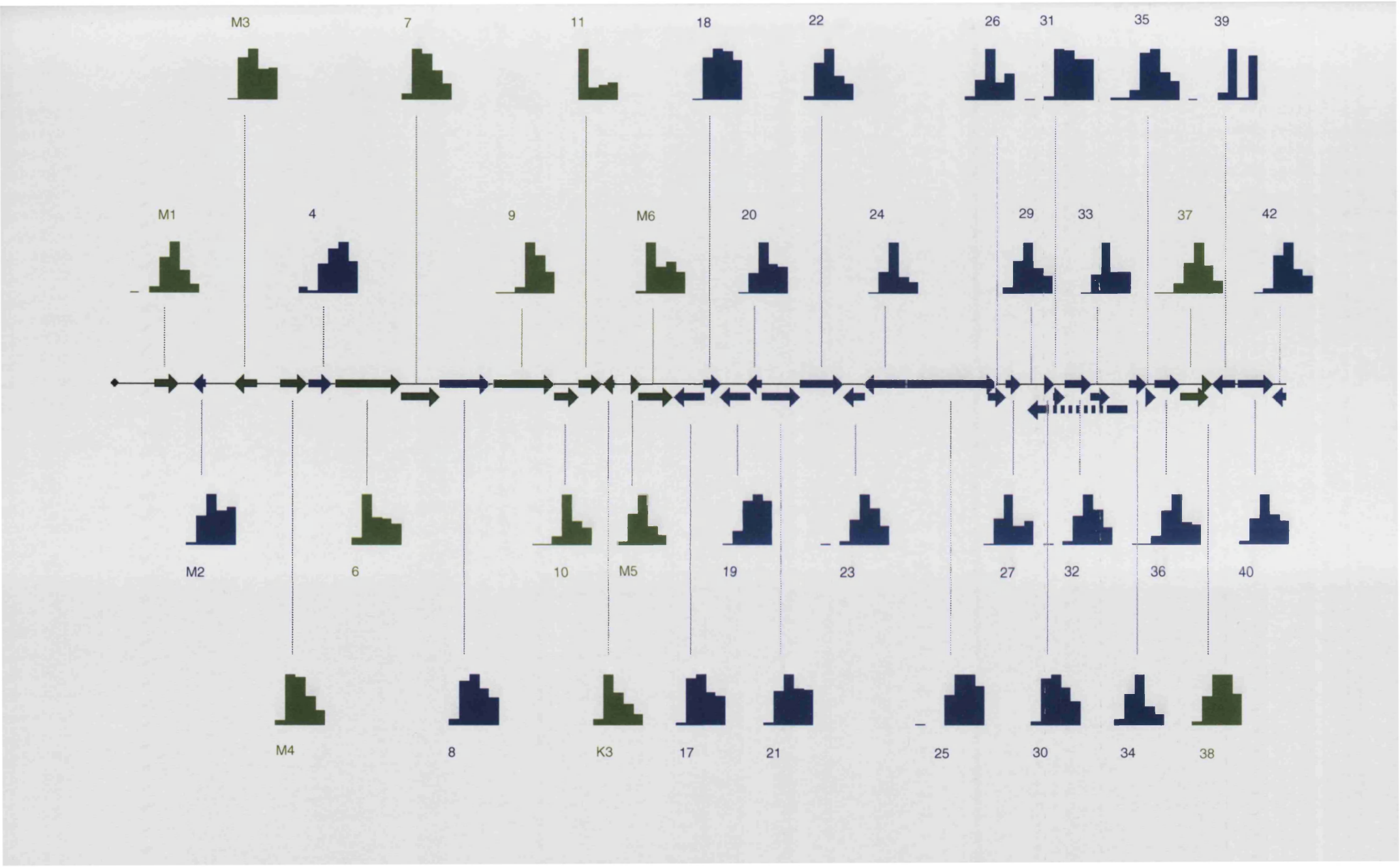
The data from the experiments presented here can be summarised as a map of MHV-68 transcriptional profiles (Fig. 8). Red arrows correspond to α-genes, defined as genes that are transcribed when *de novo* protein synthesis is blocked. Blue arrows show γ-genes, defined as genes whose expression is reduced if viral DNA replication is inhibited. The remaining genes are shown with green arrows and most likely show β-class kinetics, although this is as yet not a defined group. It should be noted that the signals generated for ORF 11 and 39 (Figure 5.2 - Figure 5.7) are very low and most likely at the threshold of sensitivity for the array. Therefore, further studies will be required to fully characterise the transcriptional profiles of these genes. Interestingly, with the exception of these two genes, it seems unlikely that any strictly latent genes are represented on the array as all the other genes show expression during a lytic infection.

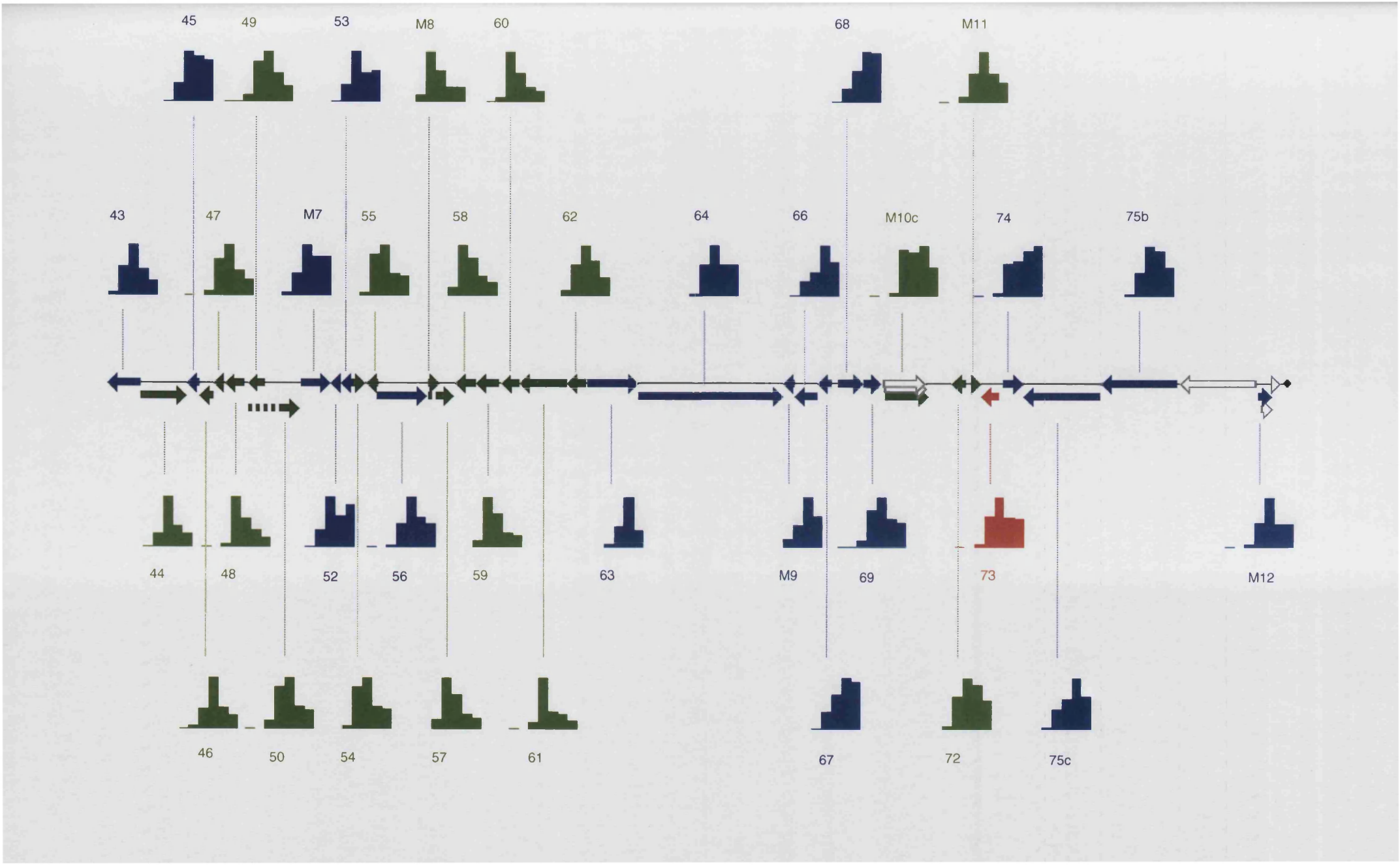
It is clear that genes with similar kinetics tend to cluster on the genome. The expression of γ-genes tends to follow one of two profiles. These are a relatively late initiation of transcription followed by a sustained or a reduced expression. However, there is a continuous range of profiles between these two extremes suggesting that there is a range of regulatory mechanisms controlling the expression of these genes.

The remaining set of genes all show early peaks followed by reduced expression, which is characteristic of β-genes. However, it has been shown here that this is also

a feature of α-genes. Although the time course data shows genes to have distinct transcriptional profiles, there is an overlap between genes of different kinetic classes. This demonstrates the need to block β- or γ-transcription to elucidate the kinetic organisation of the genome.

Figure 5.17 (on following pages). Transcriptional Map of MHV-68.
Expression profiles for the genes of MHV-68 were placed onto a physical map of the viral genome. Each expression profile is shown as a bar chart indicating transcript levels at 0h, 1h, 3h, 5h, 8h, 12h and 18h pi. Each gene's expression profile is plotted on a relative scaled, normalised to its own maximal expression. Each gene is colour coded with respect to its expression kinetics. Red arrows represent α genes, blue arrows represent γ genes and green arrows represent the remaining set of genes, predicted to be β genes. White arrows represent genes not included on the MHV-68 array.


M3
7
11
18
22
26
31
35
39
M1
4
9
M6
20
24
29
33
37
42
M2
6
10
M5
19
23
27
32
36
40
M4
8
K3
17
21
25
30
34
38


45
49
53
M8
60
68
M11
43
47
M7
55
58
62
64
66
M10c
74
75b
44
48
52
56
59
63
M9
69
73
M12
46
50
54
57
61
67
72
75c

5.6 Discussion

The aim of the work in this chapter was to use the cDNA membrane-based array (developed as described in Chapter 4) to perform a global study of MHV-68 gene expression. The work presented here includes the first comprehensive analysis of gammaherpesvirus transcription throughout a *de novo* productive infection cycle.

5.6.1 Effects of MHV-68 Infection on Host Gene Expression

Similar to other herpesviruses (Nishioka, 1977; Fenwick, 1990), MHV-68 appears to effect a rapid shutoff of host protein synthesis. Results of the differential display analysis of MHV-68 infection of BHK cells (3.4.3) showed that the poly-A RNA population of infected cells was far less complex compared to mock-infected controls. This suggested that the range of mRNA species being produced by infected cells was less than in uninfected cells. Array analysis of MHV-68 infection has confirmed this initial observation, as the abundance of host gene transcripts was observed to fall following infection of NIH 3T3 cells by MHV-68. This reduction in host RNA levels was observed as early as 1h pi (5.1.2), and continued up to 18h pi (5.1.7). This was conclusively confirmed by northern blot analysis of β-actin transcript levels following infection (5.4.4).

5.6.2 Kinetics of MHV-68 Gene Expression

The kinetics of MHV-68 gene transcription were examined by inhibiting infections with CX or 4'-S-EtdU. Only 1 viral transcript was found to exhibit α kinetics: ORF 73 transcript levels increased when *de novo* protein synthesis was blocked, relative to uninhibited infections. ORF 73 is the only herpesvirus α gene to have homologues in every sequenced mammalian and avian herpesvirus (UL54 in HSV-1). The gamma-2 herpesvirus ORF 73 homologues only conserve the C-terminal domain of ORF 73 (Albrecht, 1992; Russo, 1996; Virgin, 1997). This conserved C-terminal domain also shares secondary structure homology with the DNA binding domain of EBV nuclear antigen 1 (EBNA-1) (Grundhoff, 2003), which suggests that the gene could have shared functions across both gamma-1 and gamma-2 herpesvirus subgroups.

HHV-8 ORF 73 or latency-associated nuclear antigen (LANA) is expressed during latency (Gao, 1996a; Gao, 1996b; Rainbow, 1997) and known to interact with P-53 (Friborg, Jr., 1999) and RING3 (Platt, 1999), but it not known whether these functions map to the conserved or HHV-8 specific regions of ORF 73. As LANA is much larger than ORF 73 of MHV-68, bovine herpesvirus-4, HVS and macaque rhadinovirus, these functions could be specific to HHV-8. Therefore, the exact

function of ORF 73 in MHV-68 is unclear, but has been confirmed to be an α gene in this study.

Although only 1 α gene was identified by the array, it is possible that other genes were expressed in the absence of *de novo* protein synthesis, but below the detection sensitivity of the array. Therefore, RT-PCR was also employed to examine α-gene expression (data not shown). ORF 73 was readily detected by RT PCR in the presence of CX, which confirmed the results of the array. However, faint bands corresponding to ORFs 50, 57 and K3 could also be detected indicating that these too could be α-genes (data not shown).

Other studies of MHV-68 α-gene expression using different techniques and experimental conditions to those used here, have suggested that ORFs 57, 50 and K3 are expressed in the absence of protein synthesis (Liu, 2000; Rochford, 2001). Furthermore HSV-1 is known to express 5 α genes (Roizman, 2001a) and HHV-8 encodes more than 1 (Talbot, 1999; Zhu, 1999) so it seems likely that MHV-68 also encodes more than 1 α gene. Therefore, a more sensitive, dedicated study of α-transcription may be required to fully elucidate α gene transcription in MHV-68. Interesting there have been problems elucidating α transcription in HHV-8 (Renne, 1998; Sun, 1999). However, this could be due to having to rely on a chemically induced reactivation of the virus from latently-infected cell, instead of following a *de novo* infection.

The genes designated as β genes in this study are not a defined group, but rather classified by a process of elimination. They consist of those genes that did not qualify as α or γ genes (5.5). MHV-68's β genes are typical in that they consist primarily of genes involved in viral genome replication. The genes designated as γ genes, however, were defined, as those transcripts showing reduced abundances when viral DNA replication was inhibited by 4-S-EtdU (5.3). Typically again, these γ genes consist mainly of genes for structural components of the virus.

Hierarchical cluster analysis of the transcriptome through an uninhibited productive infection resulted in 4 groups of transcription profiles (5.1.8). The first cluster consisted of genes whose expression peaked quickly by 5h pi and then fell quickly again. They consisted mainly of genes involved in DNA replication (e.g. ORF 9 DNA polymerase, ORF 60 & 61 ribonucleotide reductase subunits), as well as candidate α genes (ORF K3, ORF 57). The majority of MHV-68's genes were represented in cluster 2. Their expression also peaked by 5h pi, but then remained at higher levels. They included an α gene (ORF 73), DNA replication genes (e.g.

ORF 54 UTPase, ORF 56 primase protein) as well as structural proteins (e.g. ORF 8 glycoprotein B, ORF 43 capsid protein).

The transcription profiles in clusters 3 and 4 were similar as all these profiles showed a late peak in transcript abundance, followed by a maintained high level. However, cluster 3 profiles peaked by 8h pi, whereas cluster 4 profiles peaked by 12h pi. The genes in these clusters consisted of structural components of the virus (e.g. ORF 25 major capsid protein, ORF 67 tegument protein).

5.6.3 Individual Transcription Profiles

As has been touched on earlier, most of the results from the array analyses are reflected in the expression kinetics of the small subset of MHV-68 genes which have been studied by various northern blots and RNAse protection assay experiments (Mackett, 1997; Husain, 1999; Simas, 1999; Wu, 2000b; Rochford, 2001). Many of the remaining genes have homologues in other herpesviruses which have been better characterised and which allow useful comparison with the data presented here. For example herpes simplex virus 1 (HSV-1) genes UL24, UL37 and UL49a have been shown to be γ-genes (Roizman, 2001a) and their predicted homologues in MHV-68, ORFs 20, 53 and 63 have been shown here to have γ-kinetics as well. Similarly, HSV-1 genes UL2, UL12 and UL50 are β-genes and in MHV-68, ORFs 46, 37 and 54 have been shown here to be β-genes as well. This demonstrates that the array is a reliable global tool for rapid analysis of MHV-68 transcript levels under a variety of conditions during infection.

The majority of MHV-68 genes have not been characterised but most do have primary sequence homologues in other herpesviruses, of which some have been studied. It is interesting to compare the transcription patterns of MHV-68 genes with those of better characterised MHV-68 genes, and also to characterised homologues of other herpesviruses.

MHV-68 ORFs 9 and 59 share very similar transcription profiles (5.1); their 2 profiles cluster with a correlation coefficient of 0.993 (5.1.8). This suggests that the 2 genes are transcribed together (they peak in transcript abundance by 5h pi after which their levels fall, and they form part of the first cluster of genes to peak during MHV-68 infection). These genes are homologous to ORFs 9 and 59 of other gammaherpesviruses, and are therefore predicted to encode MHV-68 DNA polymerase and a replication protein, respectively (Virgin, 1997). The homologues of MHV-68 ORF 9 and 59 in HSV-1 are UL30 (DNA polymerase) and UL42 (processivity factor), respectively. These 2 genes form the HSV-1 DNA polymerase

heterodimer and are transcribed together (Crute, 1989a). The array data suggests that MHV-68 DNA polymerase also consists of an ORF 9 and 59 heterodimer, as it shows that these 2 genes are likely to be transcribed together. ORFs 9 and 59 of HHV-8 also encode the DNA polymerase and processivity factor, respectively (Chan, 2000). However, HHV-8 array analysis suggests that these genes are not transcribed together by HHV-8 (Paulose-Murphy, 2001; Jenner, 2001), although these studies model reactivating virus as opposed to *de novo* infection. Further studies of MHV-68 DNA polymerase are now required to elucidate transcription of these genes by MHV-68.

Another gene, ORF 28, is found in other gammaherpesviruses, but missing from the MHV-68 genome. A probe for the inter-genic region between ORFs 27 and 29 (probe 27/29b) had been included on the array, as were all undesignated sequences of more than 100bp (4.2.2). Analysis of the array data suggested that this region encoded a gene, as a distinct transcription profile was observed for probe 27/29b. This suggested that a gene had been overlooked by Virgin *et al.*'s analysis of the genome (1997). MHV-68 was subsequently resequenced and the genome reanalysed by Milligan *et al.* (1998), and this analysis predicted an ORF 28 between ORFs 27 and 29 corresponding to probe 27/29b. The array data suggests that Milligan *et al.*'s analysis (1998) is correct.

HSV-1 is the best-studied herpesvirus and therefore provides many useful comparisons for the array data. Just 7 of HSV-1's genes are sufficient to replicate origin-containing plasmids and these genes are essential for HSV-1 replication: UL5, UL8, UL9, UL29, UL30, UL42, UL52 (Schaffer, 1973; Weller, 1983; Wu, 1988; Marchetti, 1988). Homologues of these genes are found in most herpesviruses including HHV-8 (ORFs 6, 9, 40, 41, 44, 56, 59) (Russo, 1996). However, MHV-68 has only been found to encode 6 of these and lacks an ORF 41 homologue (Virgin, 1997). It remains to be seen whether this gene has been overlooked in the analysis of the MHV-68 genome. The MHV-68 array shows that ORFs 6, 9 and 59 cluster closely together, as do ORFs 40, 44 and 56. Also clustering with ORFs 6, 9 and 59 is one gene that has no predicted function, ORF M6. It would be interesting to study MHV-68 ORF M6 further to see whether it could encode an ORF 41 homologue. Although ORF M6 was predicted to be unique to MHV-68 (Virgin, 1997), so too was ORF M7. ORF M7 has been shown to the homologous to glycoprotein 150, which is well-conserved amongst herpesviruses (glycoprotein 150; Stewart, 1996). Interestingly, the second sequence analysis of MHV-68's genome suggests that M6 is not an ORF (Milligan *et al.*, 1998).

MHV-68 ORF M3 has been characterised experimentally and has been shown to encode an abundantly secreted chemokine-binding protein (van Berkel, 1999; van Berkel, 2000; Parry, 2000; Bridgeman, 2001). The 1.4kb M3 transcript has been shown to be expressed abundantly at both 12h and 24h pi (van Berkel, 1999), but not in the presence of CX. These observations have been confirmed by the MHV-68 array data presented here, and also expanded on as the array detected M3 transcripts by 3h pi. The initial study suggested that ORF M3 was transcribed with early-late or γ-1 kinetics, as reduced transcript levels were seen on northern blots when the infection was inhibited by phosphonoacetic acid (PAA). The antiviral drug 4'-S-EtdU has been shown to be a more potent inhibitor of MHV-68 DNA replication than PAA (Barnes, 1999), and array analysis of 4'-S-EtdU inhibited infections showed that M3 transcript levels were the same as in uninhibited infections. The array data also shows that M3 is expressed by 3h pi, its transcript abundance peaks at 5h pi and then remains high until late periods of infection. The data presented here suggests that M3 is likely to be a β gene.

5.6.4 Advantages and Limitations of Genome-Level Transcription Profiling

The array-based system has many advantages over the differential display system for type of study attempted here. Most significant of these is that by using a defined probeset, there is no redundancy in the data produced. Also, an array-based system is guaranteed to provide data for each probe, which allows comprehensive datasets to be collected per experiment. These features make the array-based analysis more appropriate for the study attempted here.

Although an array-based system was much better suited for the analysis of MHV-68's transcriptome, there are a number of intrinsic limitations to this technology. Essentially, arrays function via specific hybridisations between target and probe DNA sequences. However, there is always some degree of non-specific hybridisation and therefore inter-probe competition. This particularly becomes a problem when the target sequences share homology to each other. In the case of array analysis of MHV-68 transcription, this becomes manifest with the nested sets of co-terminal RNAs. The transcripts for these genes are typically encoded one after another on the genome and share a poly-adenylation signal. Therefore, each gene's transcript will have a region of shared sequence. Any probe that is homologous to a shared region will obviously bind to more than one target transcript.

For example, ORFs 52 and 53 are predicted to be co-terminal transcripts encoded on the right-hand strand of the genome, i.e. the ORF 53 transcript also contains the sequence encoding ORF 52. Resolving individual transcription profiles from such gene structures can be problematic for methods of transcription analysis. On the array, the ORF 52 probe detects both the ORF 52 and 53 transcripts, and therefore is not specific for ORF 52 alone. In fact three labelled cDNA species result from reverse transcription of these two transcripts, one deriving from the ORF 52 transcript and two from the ORF 53 transcript. Of these latter two cDNAs (primed by the ORF 52 and 53 specific primers, respectively), one will be an artefact that binds to both the ORF 52 and 53 probes. It will also be longer than the approximately 300bp of specific cDNAs and therefore will incorporate more label as well. These factors can combine to make the signals observed on the array for such gene structures unreliable indicators of transcript abundance. In nested sets of many genes, these artefact cDNAs will add unequal background signals to the array probes and therefore an alternate method of transcription analysis must be used to elucidate the levels of transcription occurring in such sets of genes.

In fact, as there are only two members in this nested set, the transcription levels for ORF 52 and 53 on the array were still representative of actual transcription, albeit with a higher background, and this was confirmed by northern analysis in this study. Northern analysis, which resolves transcription by abundance and size, allowed differentiation between the transcripts for ORF 52 and 53 as, although it suffers the same redundancy in probe specificity, multiple analyses with individual probes can resolve this issue. Therefore, northern blot analysis can be very useful for elucidating features (such as co-terminal gene sets) that are beyond the capacity of array technology. Furthermore, the potential occurrence of such genomic features requires that one is aware of the genomic and particularly, the transcriptional organisation of the study subject, both when designing arrays and interpreting the resulting data.

Bearing these limitations in mind, the MHV-68 array has been successfully used to characterise the virus' transcriptome through a productive, *de novo* infection *in vitro*.

6 Design and Development of a Oligonucleotide-Based MHV-68 Array

6.1 Oligonucleotide-Based Array Systems

Oligonucleotide-based arrays (oligo arrays) have various differing characteristics to DNA-based arrays. Typically they employ a larger number of probes permitting more detailed studies of transcription, and therefore may be a more powerful tool for the study of viral transcription. The aim of this study was to design and develop a MHV-68 oligo array, and assess the potential advantages and disadvantages of such an oligonucleotide-based system for the analysis of MHV-68 transcription. This work was performed in collaboration with the array group at Arrow Therapeutics (UK).

The main differences between the MHV-68 DNA array and the MHV-68 oligo array systems are presented in Table 6.1. The MHV-68 oligo array was designed to be comparable to the DNA array, but also to realise the potential of the oligo array technology.

Table 6.1. Main features of DNA- and oligonucleotide-based arrays

DNA Array	Oligonucleotide array
1 probe per gene	1 probe per 100bp
dsDNA probes, around 300bp	Oligonucleotide probes, 60 bp
2 replicates of each probe	5 replicates of each probe
Probes synthesised by PCR	Probes synthesised by phosphoramidite chemistry
Probes spotted onto nylon membrane	Probes synthesised *in situ* on glass slides
Radiolabelled target	Fluorescently labelled target
Single labelling only	Single or dual labelling possible
Signals quantified using phosphoimager	Signals quantified by high resolution laser scanner

6.2 Oligo Array Probe Design and Synthesis

The probeset for the oligo array was designed at the genome level, in comparison to the DNA array, which was designed from an ORF level. One 60-mer oligonucleotide probe was designed for each 100bp of MHV-68's genome, using Hotspots software (Arrow Therapeutics). Where possible, the oligonucleotide probes were designed to match up with the DNA probes used for the DNA array so that comparisons between the 2 platforms could be made at a later stage. Control probes for housekeeping genes, luciferase and negative controls were also designed. This step was performed by the array group at Arrow Therapeutics. As the oligonucleotide probes were designed at the genome level, the probeset included multiple probes for each predicted ORF, and also probes for the inter-genic regions. Furthermore, probes to the tRNA-like sequences at the left-hand end of the MHV-68 genome were also included.

To check the validity of probes designed by Hotspots software, all probes were aligned against the whole MHV-68 genome sequence using the FastA algorithm. A few probes were found to align to more than 1 region of the genome and were therefore discarded. Any probes that extended beyond the limits of an ORF were also discarded. Finally any probes that had complementarity to more than one ORF were also discarded. All probes that matched with the DNA array probe sequences were highlighted for later comparison to DNA array data. Probes that were incomplete matches to the DNA array probes were also noted.

In total 1696 probes were designed against the MHV-68 genome. Of these, 1438 probes were complementary to ORF or tRNA sequences, while 258 probes were complementary to inter-genic sequences. Following the FastA alignments, 125 probes were discarded (equivalent to 7.3%). Seventy six of the discarded probes were complementary to the repeat regions of MHV-68. Twenty two probes matched more than 1 region of the genome, and the remaining 27 probes corresponded to a human error when running the probe design software.

The probeset was replicated 5 times and each probe positioned randomly on the array. The finalised probe layout was used as a template for automated *in situ* synthesis of the oligomer probes onto glass slides. It should be noted that the resulting MHV-68 oligo arrays are physically much smaller than the DNA arrays, as the oligonucleotide probes are present at much higher spatial densities than the DNA probes. Unlike other oligo array systems, which tend to use soft-lithography or optical deproteination to synthesis the oligonucleotide probes on the arrays, this system uses phosphoramidite chemistry to synthesis the probes directly *in situ* on

the glass slides that support the array. Importantly this system is more flexible than other systems as the reagents are delivered via an inkjet system and there is no hard template that requires to be manufactured for each probeset. This flexibility allows for rapid modification of probesets, as the array can be altered as quickly as new probe sequences can be designed.

As a further control for the oligo arrays, an internal control probe was included. This control probe was used to form a grid on the array to aid the orientation of the probes, and to show that hybridisation had occurred successfully. A pre-labelled oligonucleotide target (incorporating a green fluorescent tag) complementary to the control probe was added to the pool of cDNA target prior to hybridisation. The target cDNAs were labelled with a Cy5 fluorescent tag, allowing simple differentiation between the internal control probes and MHV-68 probes.

6.3 Oligo Array Analysis of MHV-68 Transcription

6.3.1 RNA Template

RNA samples that had previously been used for analysis with the DNA array were used again for analysis with the oligo array. The RNA was isolated from NIH 3T3 cells infected at moi 10, as well as from mock-infected controls. Although RNA samples were stored at -80°C, they were checked again to ensure that no degradation had occurred using a Bioanalyser 2001 (Agilent Technologies).

Each RNA sample was analysed at 2 concentrations. The electropherogram for every sample showed sharp, well defined peaks, with a higher peak for the 28S rRNA than for the 18S rRNA. These features indicated that there was minimal degradation of the RNA samples. Furthermore, the electropherograms also showed that equivalent quantities of RNA were present in each sample, confirming the spectrophotometry-based quantification made previously. The electropherograms of RNA samples isolated at various times pi are shown overlaid in Figure 6.1. This analysis showed that the RNA samples were suitable for use as template in the oligo array experiments.

6.3.2 Target cDNA Synthesis

The RNA samples were used to synthesis fluorescently-labelled cDNA target in the same way as has been employed for the DNA array system, except that the incubation time was increased to 16h (a shorter incubation time is used in the DNA array system as RT is denatured over time by the radioactive decay of the labelling isotope).

The probes on the DNA array corresponded to 300bp fragments at the 5' end of each ORF. The specific primers used to prime the cDNA target labelling reaction were designed to only reverse transcribe the region of mRNA complementary to the DNA probe sequences. However, the probes on the oligo array corresponded to the entire genome. Therefore, oligo d(T) or random nonomers would need to be used in order to make use of all the probes on the array. As some of the probes on the array corresponded to tRNA-like sequences, which were unlikely to posses a polyA tail, random nonomers were chosen to prime the oligo array target synthesis reaction.

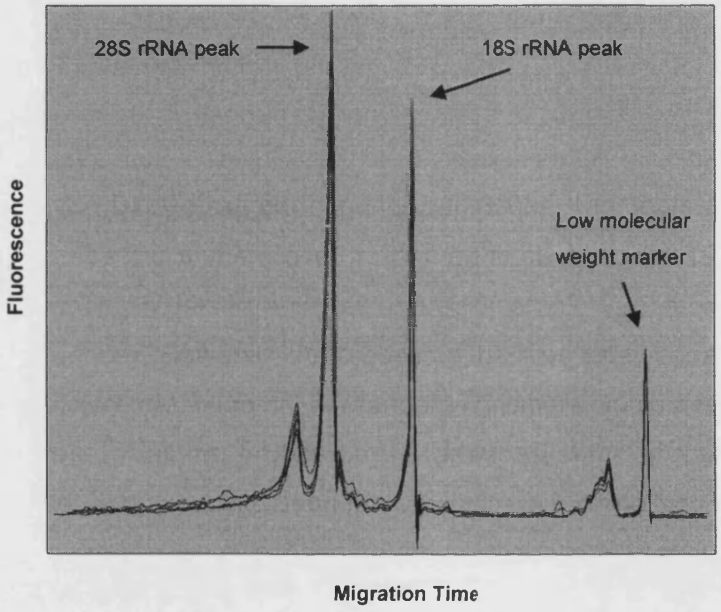


Figure 6.1. Analysis of RNA integrity.
RNA samples isolated at 5, 8, 12 and 18h pi, as well as from mock-infected samples, were fluorescently labelled and electrophoresed through a RNA LabChip (Bioanalyser, Agilent). The resulting electropherograms are shown overlaid to allow simple comparison of the profiles. The y-axes show the fluorescence and the x-axes show the migration time of RNA through the gel matrix of the RNA LabChip. Each sample consisted of 200ng of RNA.

However, for the purposes of this initial study, specific primers were used to allow better comparison to the DNA array system. Hybridisations performed with specifically primed target synthesis were found to produce weaker signals overall compared to random nonomer primed target synthesis (Figure 6.2). However, the dynamic range of signals was similar, whichever primers were used.

6.3.3 Oligo Array Hybridisation

The target cDNA was hybridised to the glass oligo arrays for 16h, following standardised protocols (Arrow Therapeutics). After hybridisation, the arrays were washed and then dipped in ether to dehydrate the array, before being air dried.

6.3.4 Quantification of Oligo Array Data

After scanning the arrays, it became obvious that the several signals on the array were beyond the range of the scanner. Therefore, a series of calibration scans were performed using varying laser strengths (10% - 100% of maximum, in 10% steps). Repetitive scanning of arrays had previously been found to have no effect on the strength of fluorescently-labelled target up to 20 consecutive scans (Dr Pete Corish, personal communication). RNA isolated at 12h pi was used for these calibrations as the DNA array had shown maximal signal strengths at this time point. All signals on the array were found to be within the range of the scanner when the laser strength was reduced to 20%. Therefore, all subsequent arrays were scanned in this way.

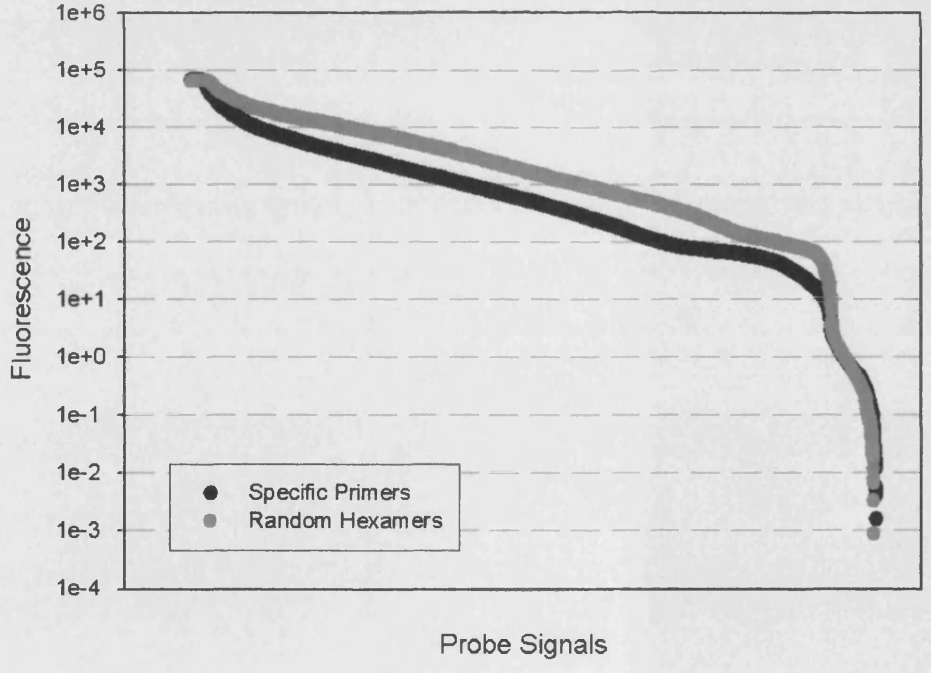


Figure 6.2. Range of signals on the oligo array.
The probe signals on the oligo array were ordered by magnitude and plotted on a scatter graph. The signals from 2 array hybridisations, identical except for the method of priming cDNA target synthesis, are shown. cDNA synthesis was either primed by random nonomers or viral mRNA specific primers (listed in Appendix I) as indicated in the key.

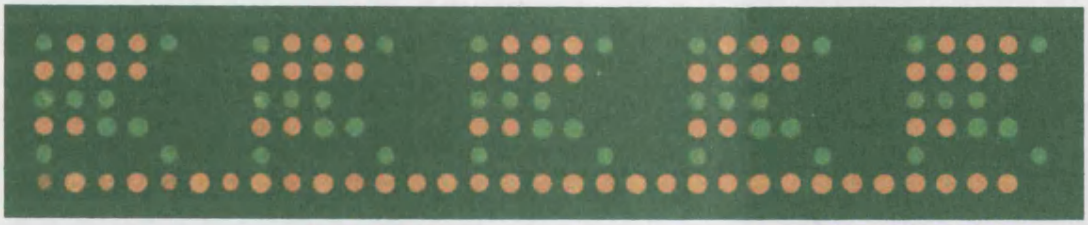
To estimate the range of the scanner, all probe signals were plotted in order of intensity (Figure 6.2). The plot shows that there is a linear distribution of signals over 7 x 10^2 fluorescence units, which was similar to the dynamic range of the DNA array. The lower limit of the scanner was set as 100 units, based on this analysis.

In addition to the signal intensity for each probe, various other criteria were also calculated and recorded. In total 50 properties were calculated for each probe, including: Position on the array (to aid identification of each probe), error measurements for each signal, local background for each probe, uniformity of signal across a spot, and whether the signal was saturated. As there were 5 randomly placed replicates of each probe and many probes per gene, the dataset was considerable in size and allowed increased confidence in the data from a single array, in comparison to the DNA array. The dataset from each hybridisation was imported into Excel software for further analysis.

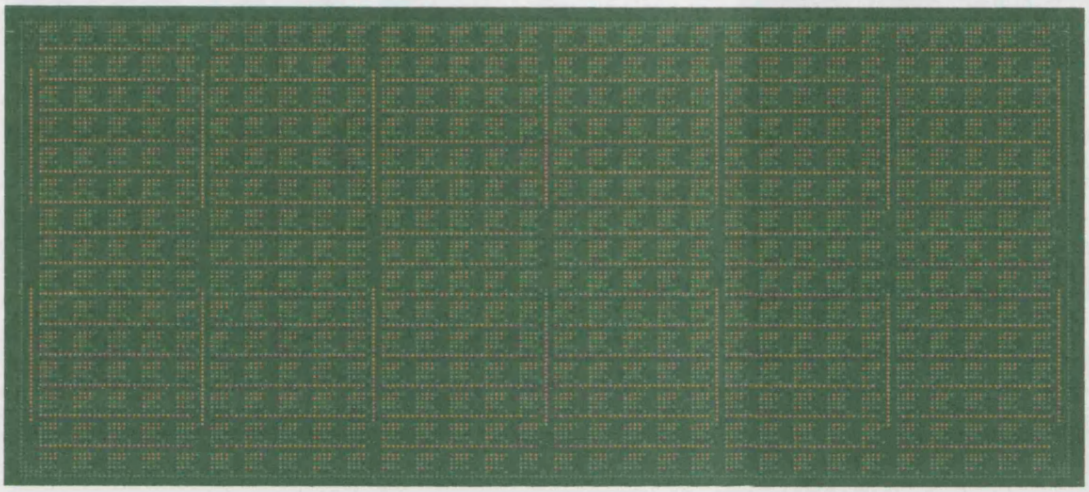
An example of a scanned array image is shown in Figure 6.3. This test array consisted of 2 probes only. These 2 probes were positioned on the array in a repeating series so that any inconsistencies would be obvious following hybridisation. The target mix consisted of Cy3-labelled 25mer oligomer, Cy5-labelled pyrimidine-only control oligomer, Cy5-labelled 60mer 'long' oligomer target (all oligomer targets at 0.2nM), a Cy5-labelled bacterial gene cDNA target and a Cy3-labelled GFP cDNA target (both cDNA targets at 10ng). The Cy3-labelled 25mer oligomer and the Cy5-labelled 60mer 'long' oligomer were complementary to the probes on the array, and the other target species were used to simulate non-specific background hybridisation to probes. This test array showed that there was little inconsistency in the hybridisation of the control oligomers.

Figure 6.3 (following page). Scanned image of hybridisation to a test array.
The test array consisted of a repeated series of 2 probes. The target pool consisted of fluorescently labelled oligomers complementary to the probes and various background species. A. Close up image of a small section of the array. Individual spots are shown which bound the Cy3-labelled target (green), the Cy5-labelled target (red), or no target. B. Zoomed out image to show the entire array. The effect of the repeated series of probes on the array can be seen. Also probes used to orientate the array can also be seen, such as those marking the borders of the array.

A.



B.



6.4 Interpretation of Oligo Array Data

6.4.1 Comparison of Data from Oligo Arrays and DNA Arrays

A time course experiment was set up using RNA was isolated from NIH 3T3 cells, infected with MHV-68 at moi 10, for 5h, 8h, 12h and 18h pi. Mock-infected controls were also taken. The same specific primers used for the DNA array, were used to produce fluorescently-labelled cDNA target for the oligo array. The oligo array hybridisation was performed, and the results compared against similar data from the DNA array system.

Only those probes on the oligo array that matched the probes used on the DNA array were considered. Each probe on the oligo array was replicated 5 times and in addition, each ORF of MHV-68 was represented by up to 4 different probes. For the purposes of this comparison, the mean signal for all probes corresponding to a given ORF on the oligo array was compared to the mean signal for all probes corresponding to a given ORF on the DNA array.

The mean signal for each ORF was plotted as a bar chart as shown in Figure 6.4. For comparison purposes, the equivalent data from experiments with DNA arrays are also shown. While still preliminary data, there is good general concordance between the data from the DNA array and oligo array platforms. In particular the relative abundances of one transcript relative to another at any time point are well replicated by the 2 array systems. Also when 2 DNA array hybridisations show similar profiles (e.g. at the 12h and 18h time points), so do the oligo array hybridisations. It should be noted that the DNA array data shown in Figure 6.4 represents the mean of several experiments (n = 4-6), whereas the oligo array data summarises the 5 replicate probesets on a single array.

Figure 6.4 (on following pages). Results of hybridisations to oligo arrays.
RNA was isolated from NIH 3T3 cells at various periods pi with MHV-68 at moi 10 (C., E., G., I.) and also from mock-infected controls (A.). Labelled target was synthesised from the RNA samples and hybridised to the oligo arrays. The mean signal for each ORF was calculated and plotted as a bar chart, for each time point (the results show the mean signal from 5 replicate probesets on a single array). For comparative purposes, the equivalent data produced by the MHV-68 DNA arrays, as shown previously in Figure 5.3 - Figure 5.7, was also plotted (B., D., F., H., J.).

Fig. 6.4A. Oligo array hybridisation - 0h pi

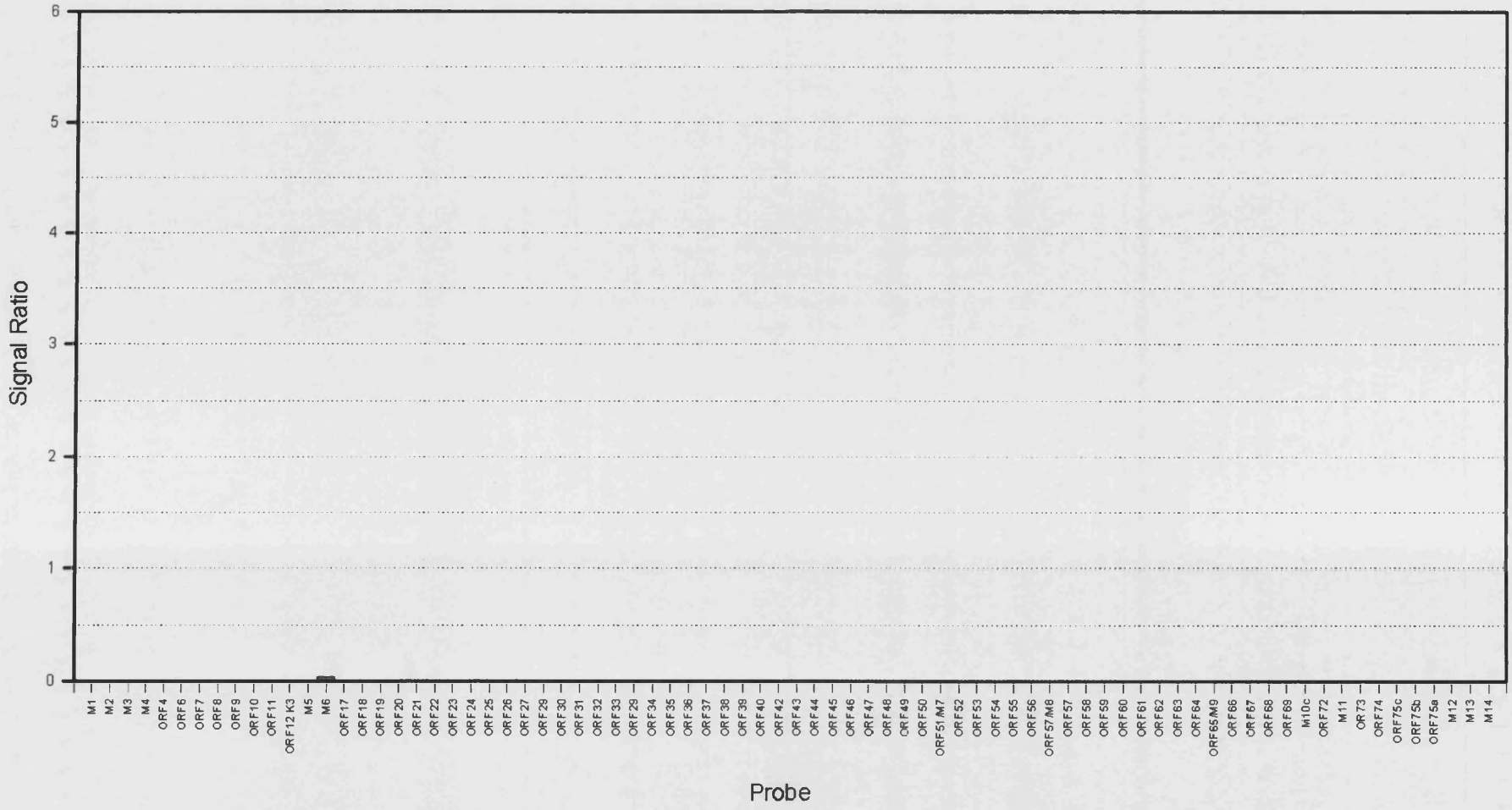


Fig. 6.4B. DNA array hybridisation - 0h pi

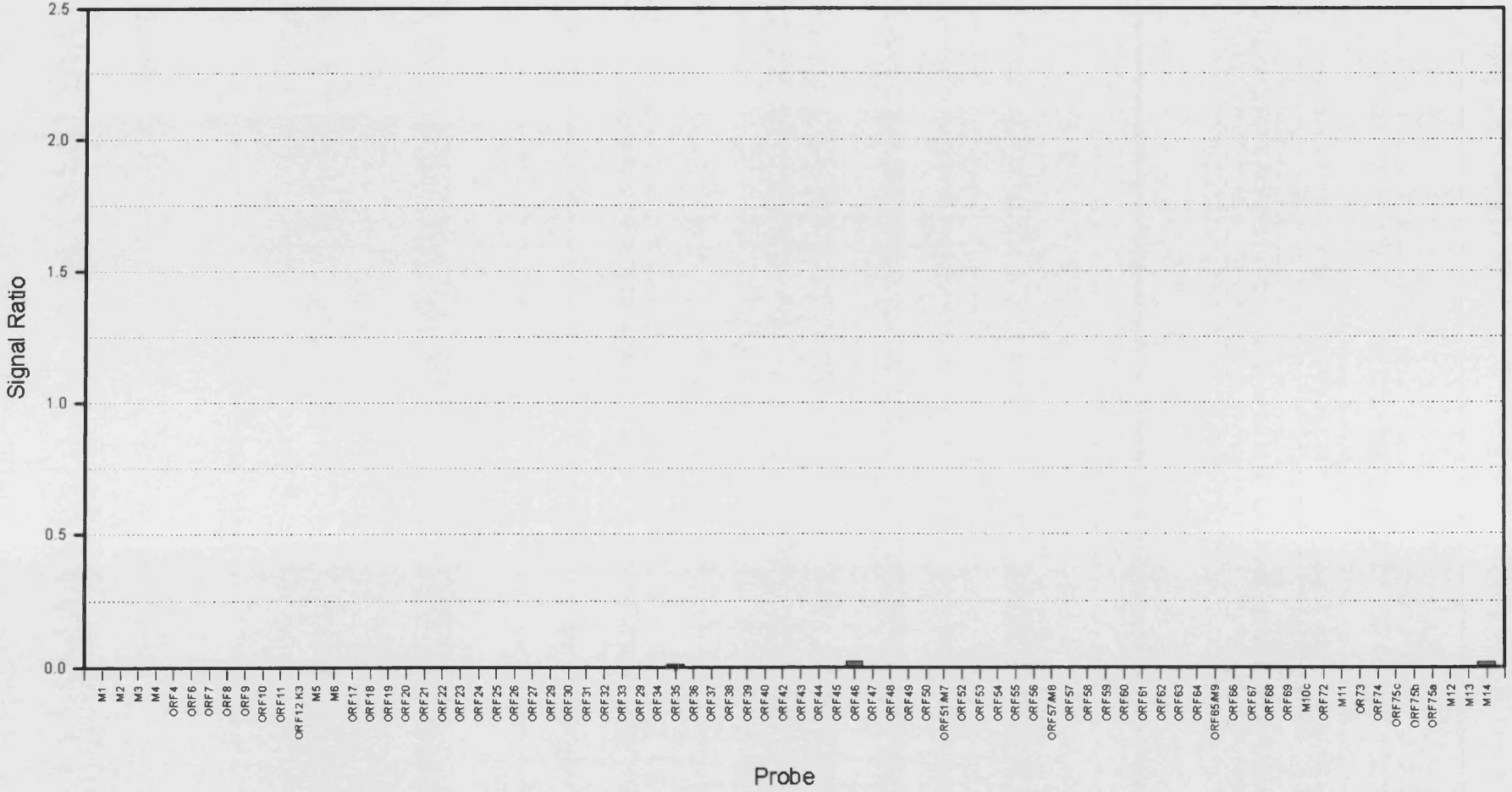


Fig. 6.4C. Oligo array hybridisation - 5h pi

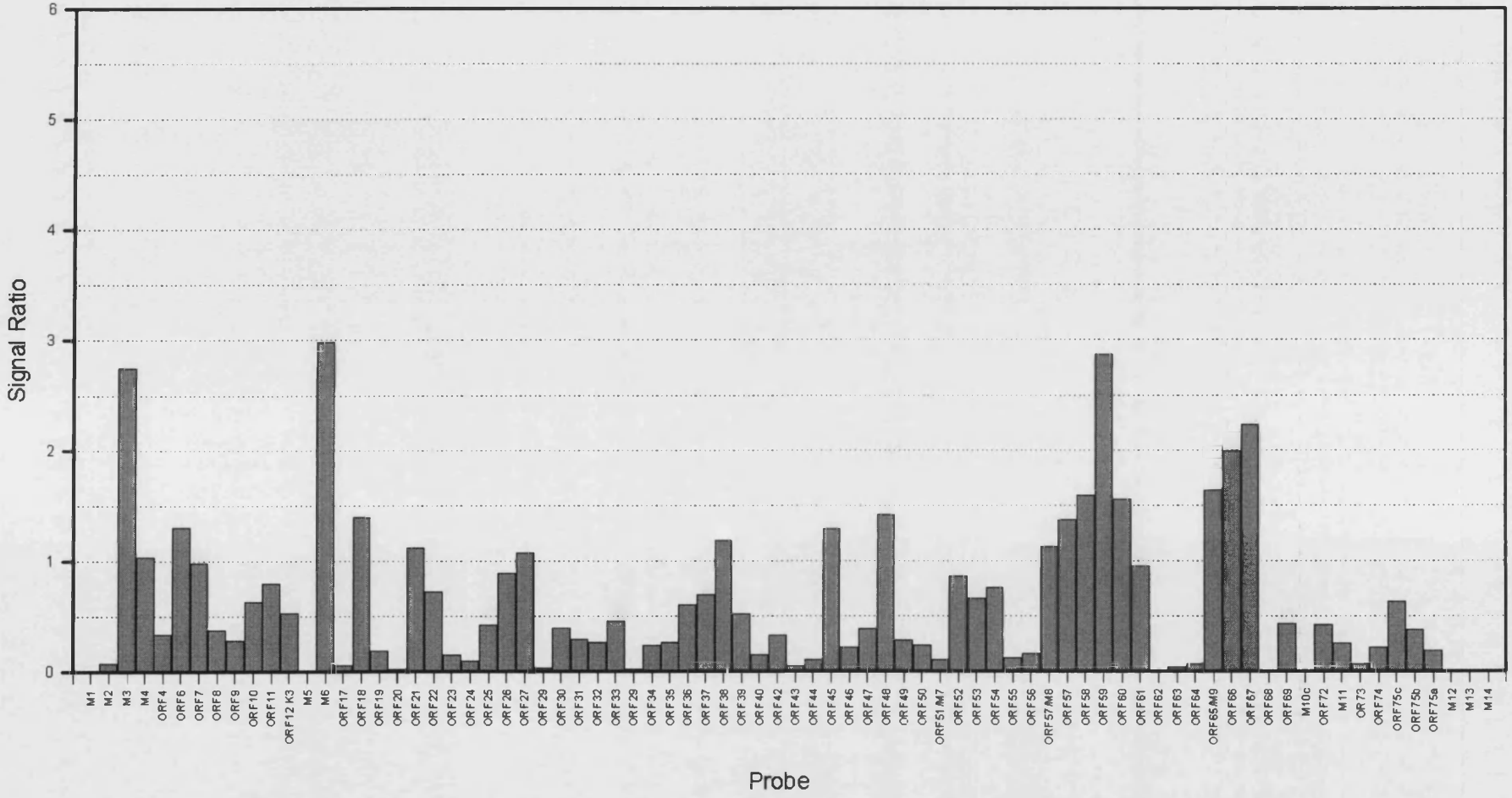


Fig. 6.4D. DNA array hybridisation - 5h pi

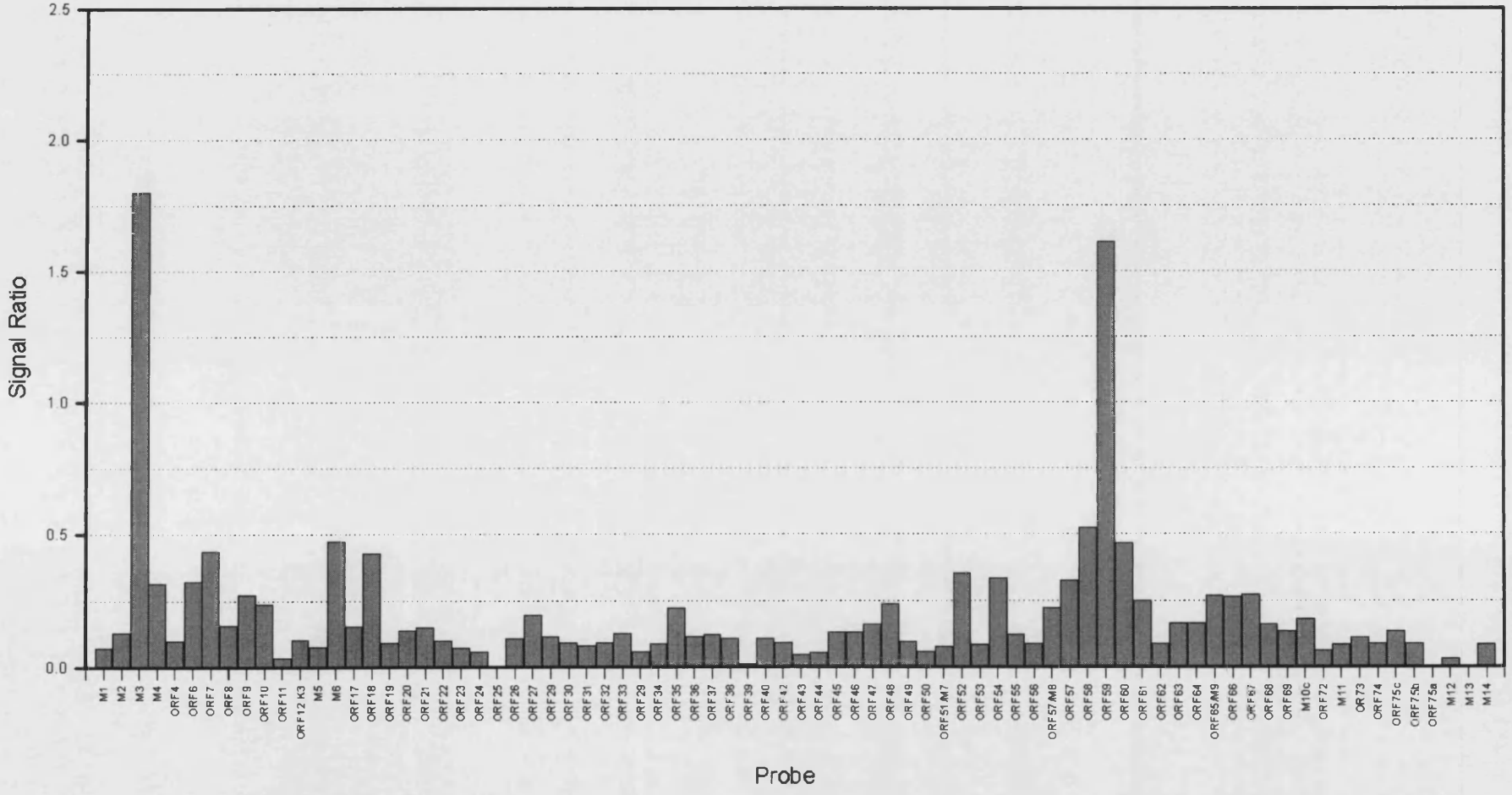


Fig. 6.4E. Oligo array hybridisation - 8h pi

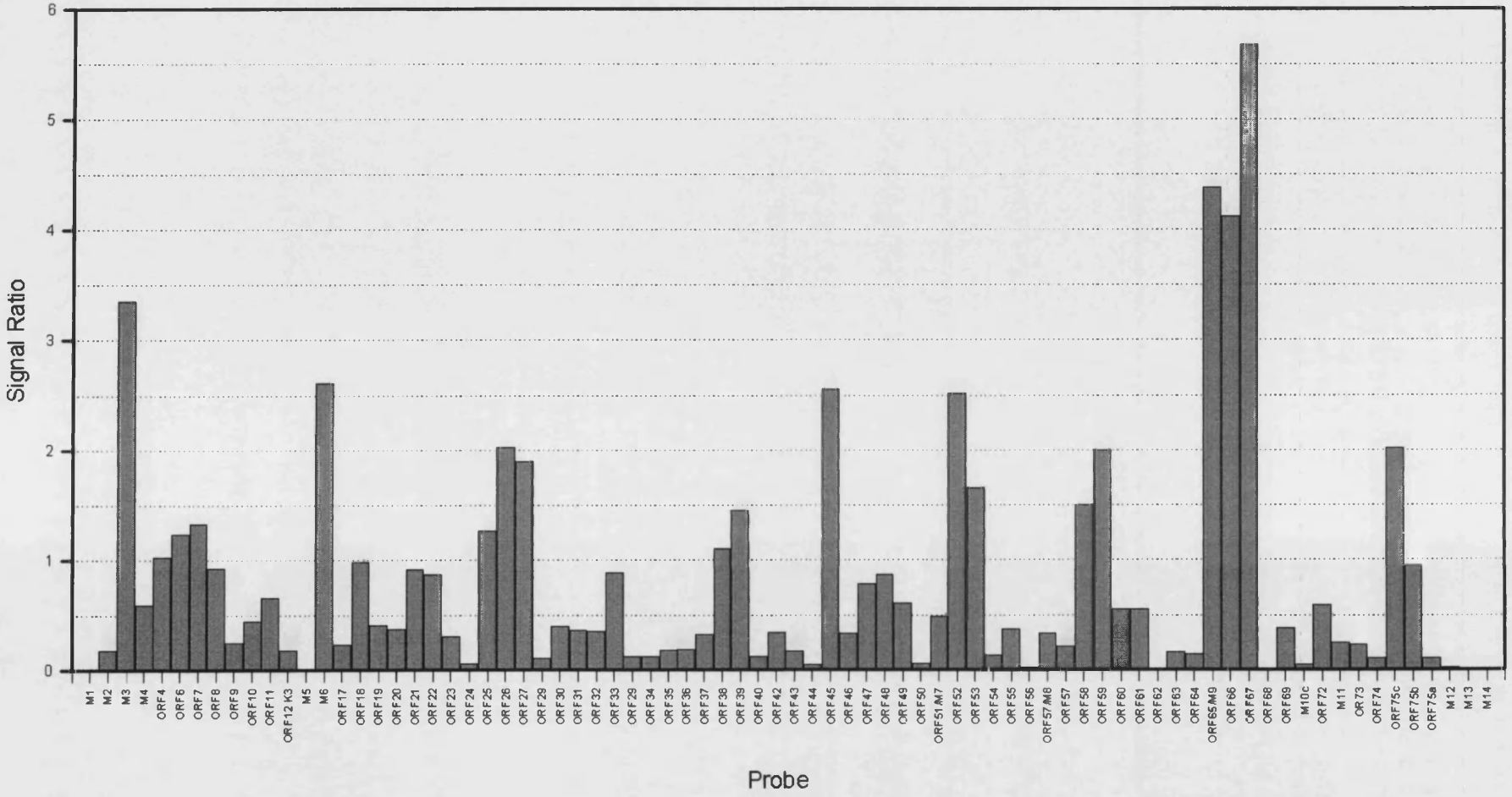


Fig. 6.4F. DNA array hybridisation - 8h pi

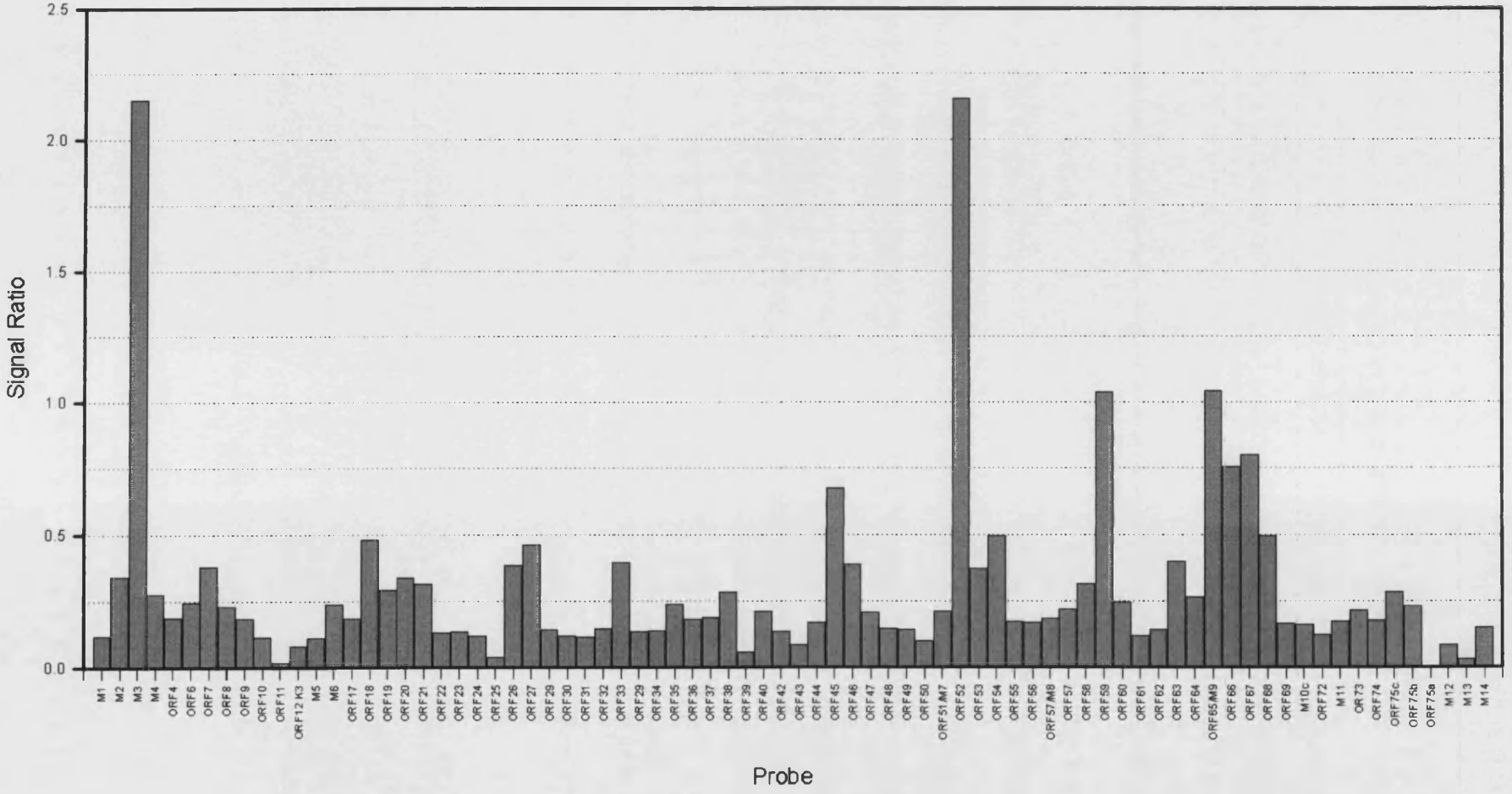


Fig. 6.4G. Oligo array hybridisation - 12h pi

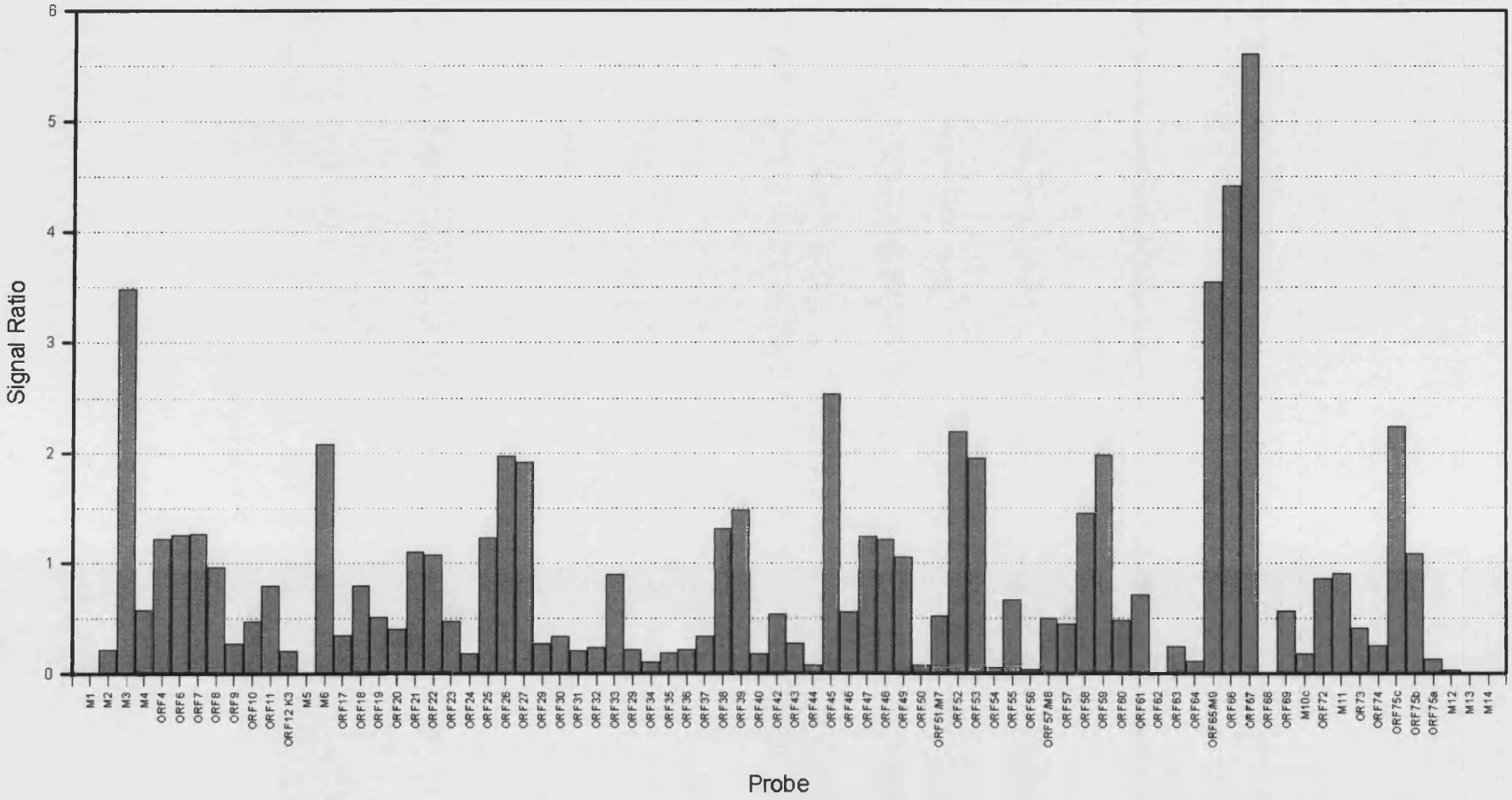


Fig. 6.4H. DNA array hybridisation – 12h pi

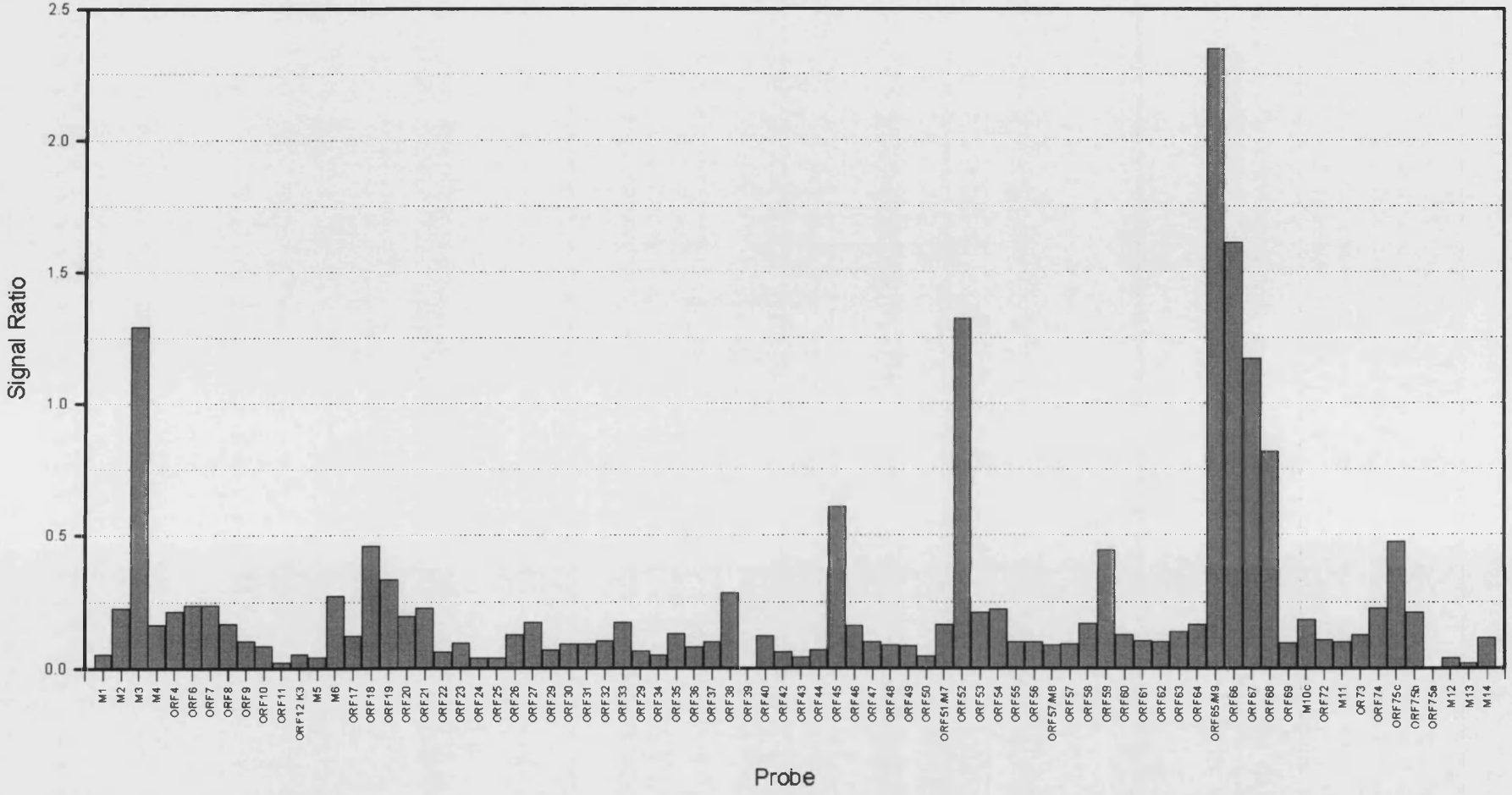


Fig. 6.4l. Oligo array hybridisation - 18h pi

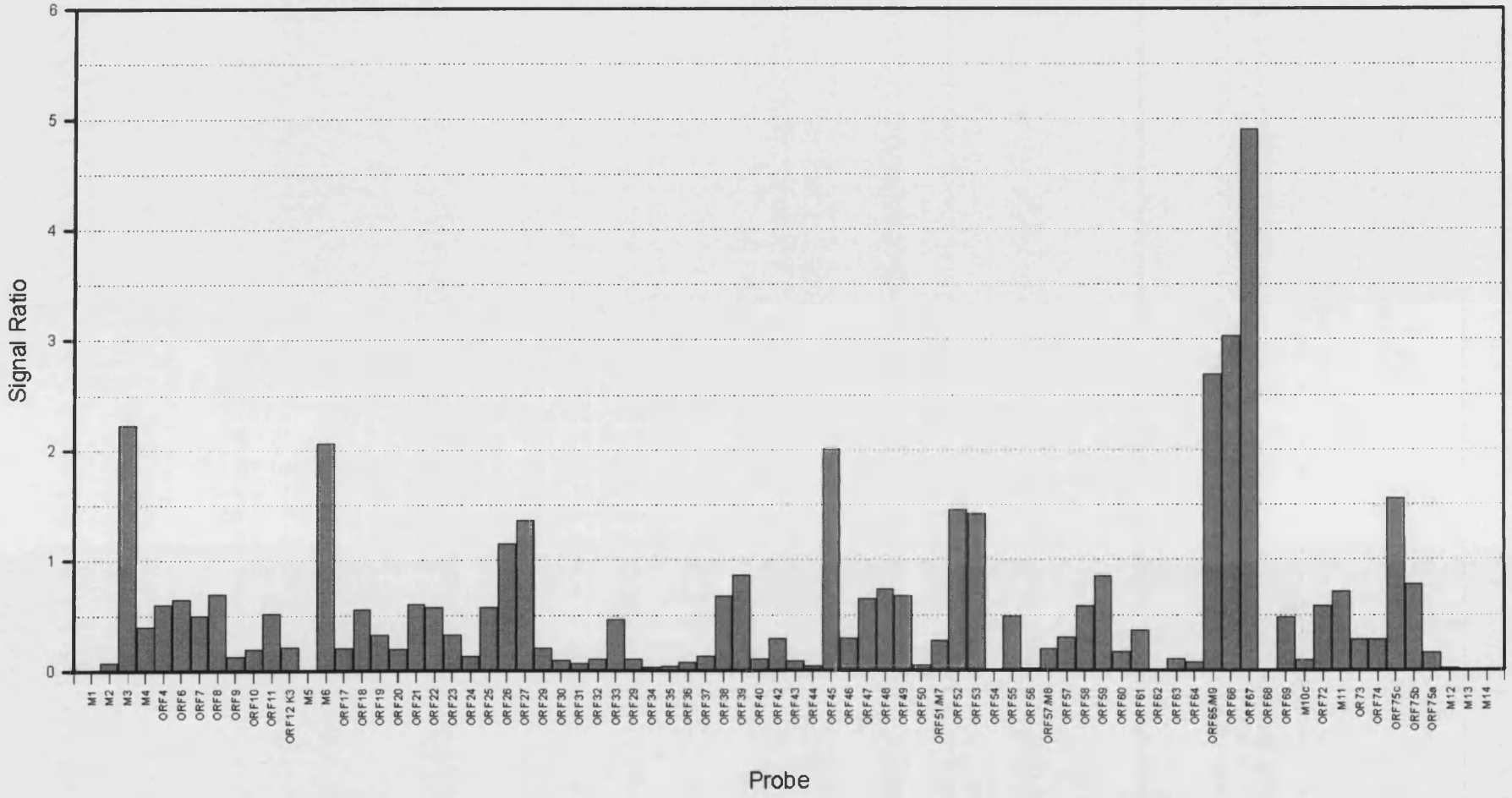
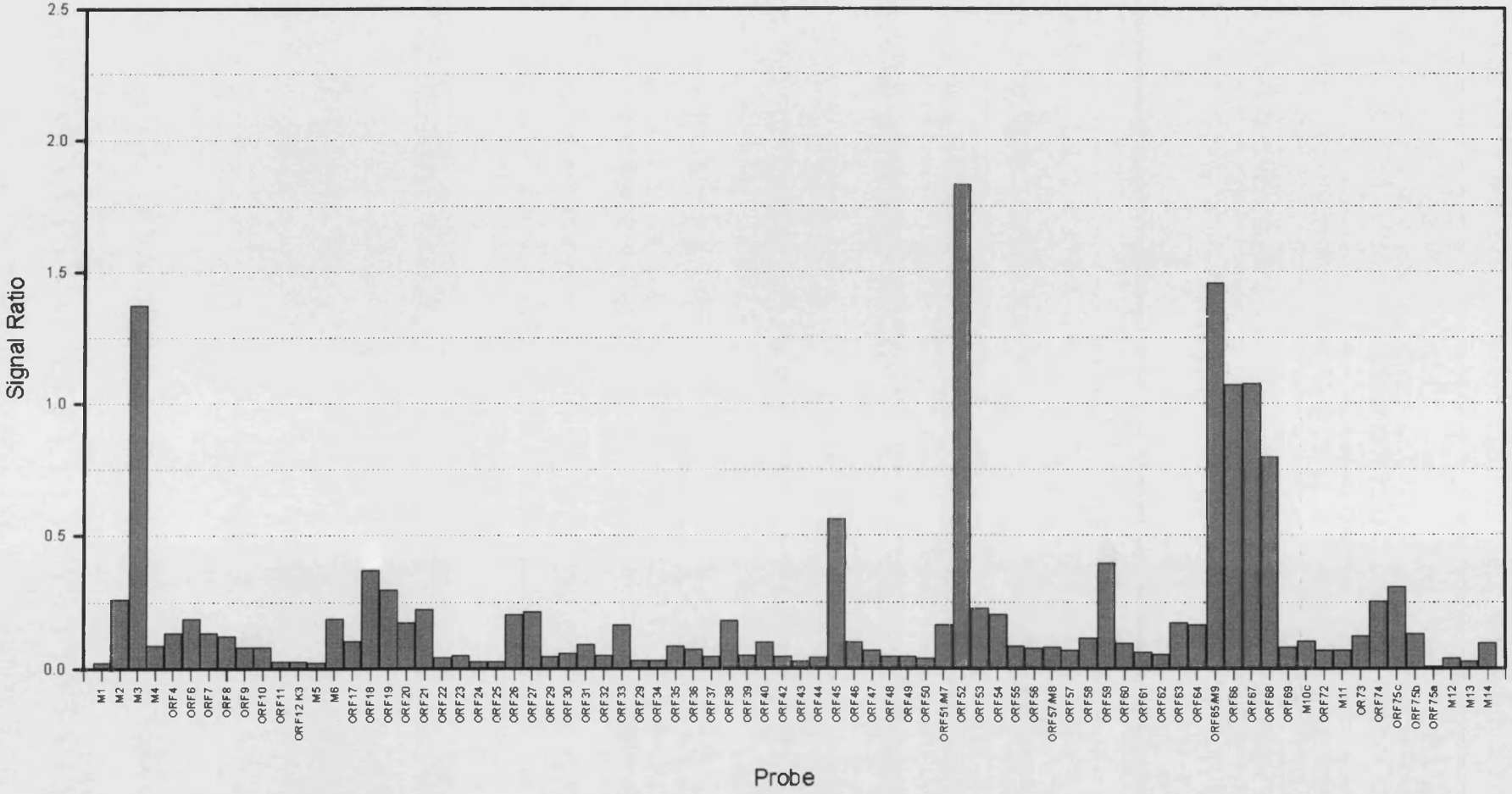


Fig. 6.4J. DNA array hybridisation - 18h pi



6.4.2 Cluster Analysis of Oligo Array Data

As well as the quantitative expression profiles, a relative cluster analysis was performed on the oligo array data. As explained previously, cluster analysis takes a gene's transcription profile, relative to itself, and compares it to all others. Therefore, it will highlight genes with similar patterns of expression, although not necessarily at similar levels.

Hierarchical clustering of the oligo array data was compared to the same analysis of the DNA array data (Figure 6.5). The overall picture of MHV-68 gene expression, as presented by the oligo array data, appears to have a higher resolution than that produced by the DNA array. This is visualised in Figure 6.5 as an increased number of shades of red. It seems likely that this increased resolution is related to the use of 60bp and 300bp probes in the oligo and DNA array systems, respectively. Presumably a physically smaller probe is able to differentiate better between similar quantities of target cDNA than larger ones.

The cluster analysis data for the oligo and DNA arrays data showed that the expression profiles of MHV-68's genes are closely related, as demonstrated by the short branches in the respective dendograms. Both analyses also show a group of genes that peak in transcript abundance at 5h pi, shown in Figure 6.6. The oligo and DNA arrays found 22 and 16 transcripts that peaked in abundance at 5h pi, respectively, of which 11 were identified by both. Interestingly, the transcripts identified by the oligo array clustered into 2 separate groups: one cluster showing much reduced transcript abundance from 8h pi, and the other cluster showing sustained transcript abundances up to 18h pi. Furthermore, the genes identified by the DNA array also consisted of these two profiles. Therefore, it seems likely that these 2 groups of genes represent 2 different programmes of expression.

An interesting difference between the two arrays is that the oligo arrays shows some transcripts are not detected at all (ORFs M1, 62, M13, M14), whereas the DNA array detected some signal for every probe. This difference could be explained if the oligo array system produces less background, or that it's calculation of background is more efficient, than that for the DNA array. If so, this suggests that the oligo array is the more sensitive system of the 2. In fact on average, background levels were 19% of the mean signal on DNA arrays, and 15% of the mean signal on oligo arrays, or alternatively, the DNA arrays produced 20% higher background levels than the oligo array.

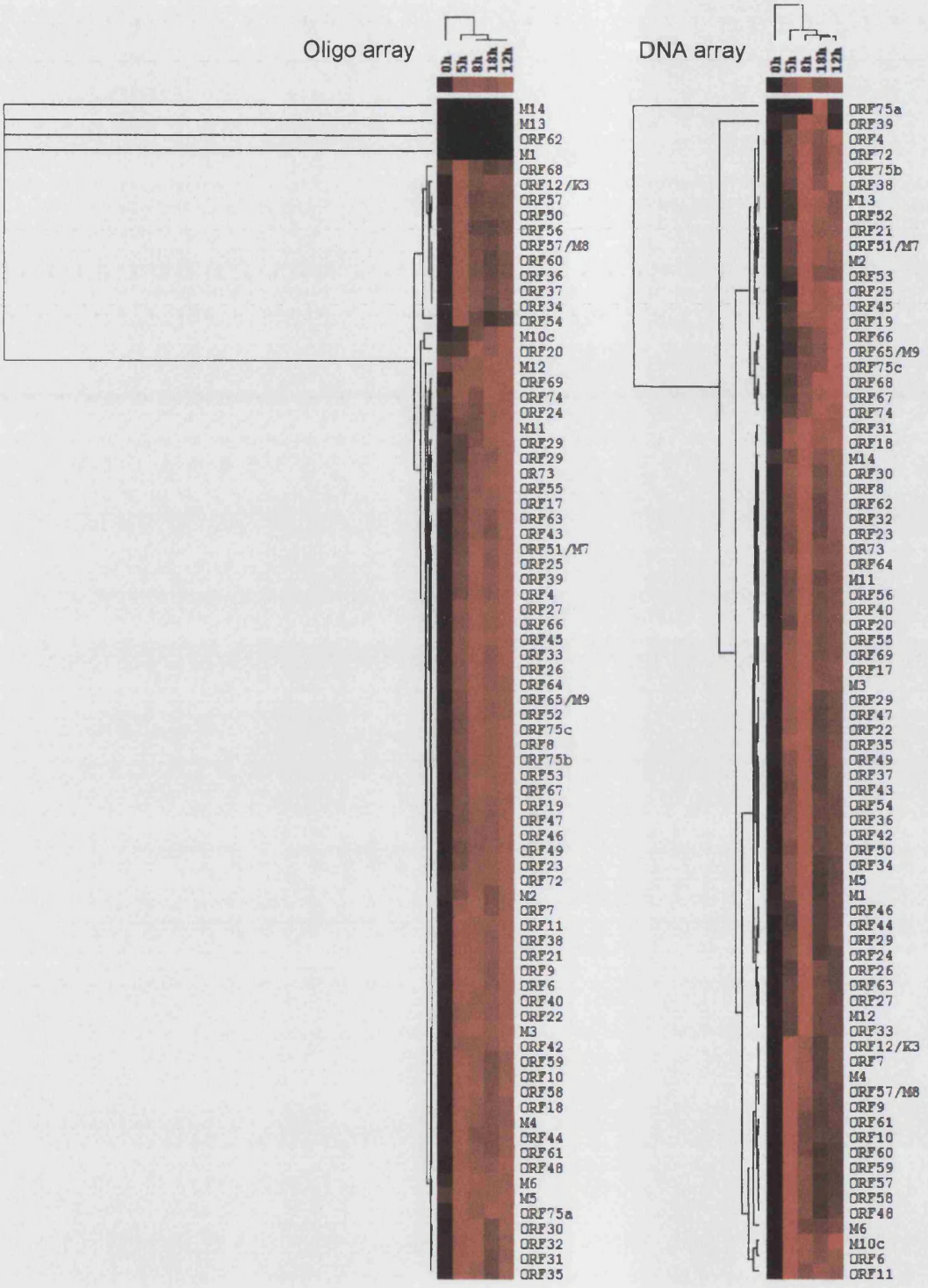


Figure 6.5. Hierarchical cluster analysis of MHV-68 expression profiles produced by the oligo and DNA array systems.

Data in the form of log2 normalised means was converted into a percentage of each gene's maximum. These percentages were imported into Cluster software and hierarchical clustering performed using the average-linkage algorithm and an uncentred correlation similarity matrix. The data is shown as a colour matrix with columns representing time points

(h pi) and rows representing each gene's expression profile. The order of the columns is not chronological but is intended to allow maximum contrast between profiles. Black boxes represent no expression and brighter shades of red correspond to increasing expression. The dendogram shows related expression profiles on the same branch, with branch lengths representing the degree of similarity between individual profiles.

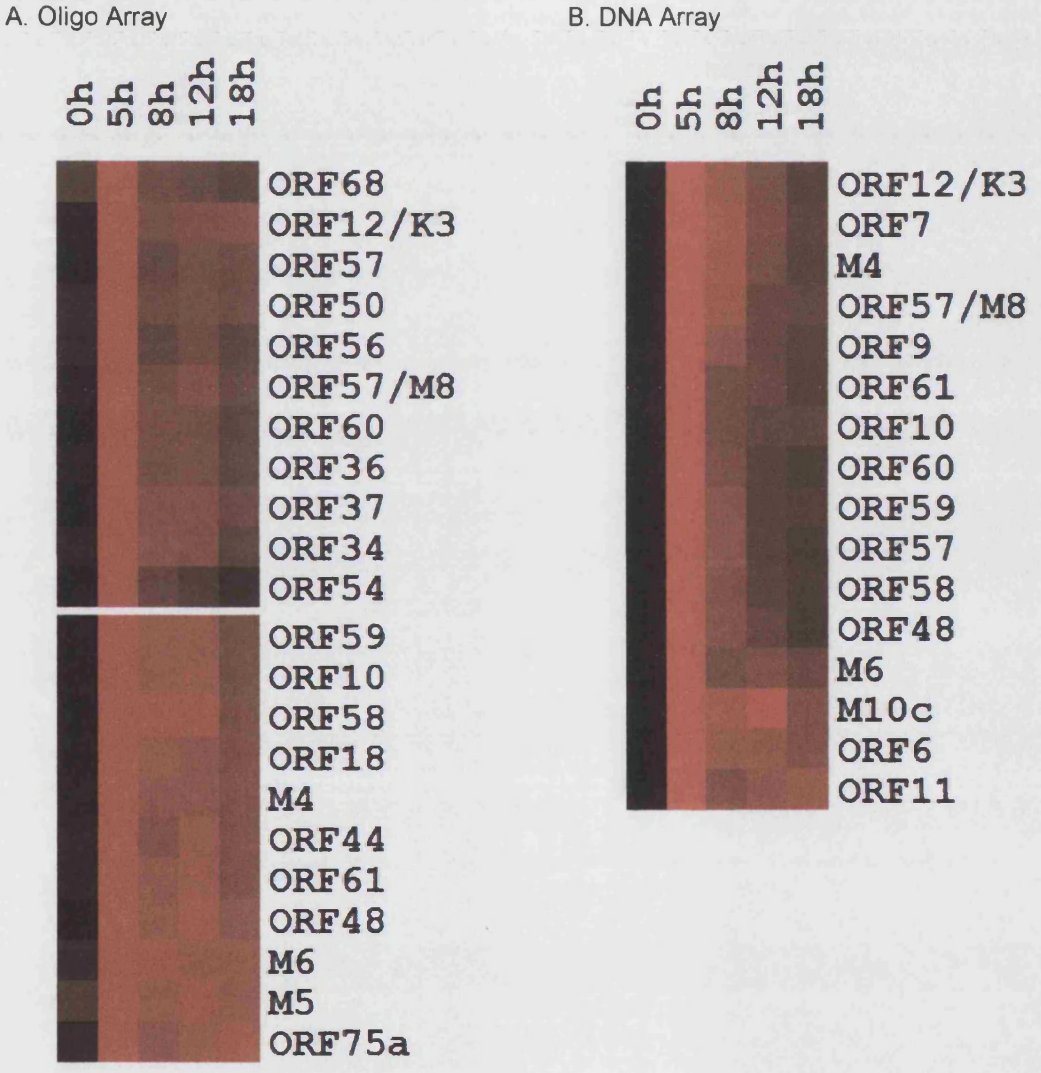


Figure 6.6. MHV-68 transcripts peaking in abundance at 5h pi as identified by oligo and DNA arrays.

Details of the analysis as are found in the legend for Figure 6.5. Only those genes that peak in transcript abundance at 5h pi are shown here.

6.4.3 Transcriptional Profile of MHV-68 tRNA Sequences

One of the advantages of the oligo array system was that its 60-mer probes could be used to detect very small transcripts. MHV-68 encodes 8 tRNA sequences which are known to be expressed during lytic and latent infection (Bowden, 1997). RNA was isolated from MHV-68 infected NIH 3T3 cells and random nonomers used to prime the target synthesis reaction. Following hybridisation and washing, the arrays were scanned and signals quantified. To analyse the array data for these sequences, the mean signal strengths for each tRNA probe were plotted as a bar chart (Figure 6.7).

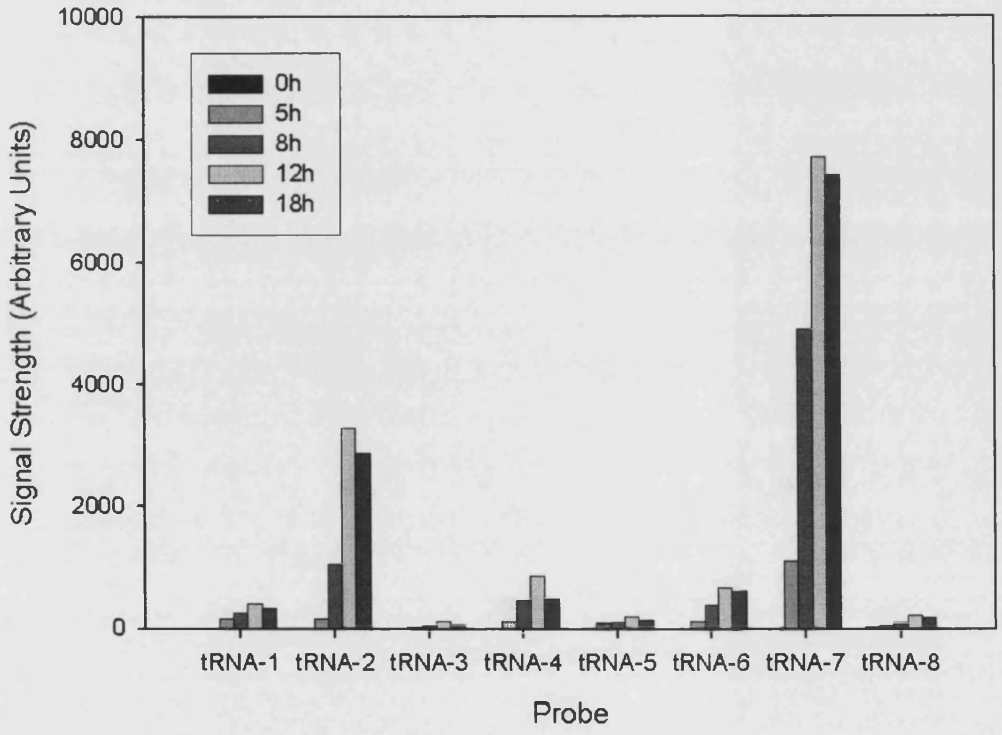


Figure 6.7. MHV-68 tRNA transcript abundances as detected by the oligo array.
NIH 3T3 cells were infected with MHV-68 (moi 10) and RNA isolated at various times pi, as well as from mock-infected controls. Fluorescently-labelled target was synthesised with random nonomer primers and hybridised to arrays. Signals were quantified and mean values for each probe calculated. The results for the tRNA sequences are shown plotted above. The key shows the time point pi that each bar represents.

Previous studies have shown that MHV-68 tRNAs are expressed abundantly during lytic infection but that tRNA3 is weakly expressed, and this is confirmed here (Bowden, 1997; Simas, 1998a; Simas, 1999). It is unclear what role these tRNAs

play, but their varied expression levels suggest that their transcription is regulated. This becomes more apparent when the organisation of the tRNAs in the genome are considered (Figure 6.8).

Figure 6.8. Organisation of MHV-68 tRNAs in the genome.
This representation of the left-hand end of the MHV-68 genome show the organisation of the 8 tRNAs and first 3 ORFs, represented by arrows. The direction of the arrows indicate the genomic strand on which the transcript is encoded.

tRNA7 is the most abundant tRNA, which suggests a functional role. However, previous analyses of MHV-68 tRNAs have omitted tRNA7 as it's sequence suggests that it would not be involved in translation (Bowden, 1997). Furthermore at least 4 of the tRNAs are not aminoacylated by host tRNA-synthases, also suggesting that the role of these tRNAs is not in translation (Bowden, 1997).

6.4.4 High Resolution Analysis of Transcript Abundance

The oligo array features 1 probe per 100bp of genomic sequence, as opposed to 1 per ORF for the DNA array. This results in a higher resolution of transcript abundance data. One possible application for this type of data is transcript mapping. Although there are several methods that are used exclusively for mapping transcripts, the oligo array can also be used to fulfil this role.

MHV-68 ORFs M8 and 57 have been predicted to encode 2 separate genes (Virgin, 1997) and 2 exons of the same gene (Milligan, 1998). Both analyses agree on the presence of a non-coding region of around 300bp between these 2 ORFs or exons. To analyse transcription of mRNA from this region of the MHV-68 genome, RNA was isolated from infected NIH 3T3 cells 12hpi, and used to synthesise labelled cDNA target to hybridise to the array. Random nonomers were used to prime the reverse transcription reaction.

For the purposes of this analysis, the probes corresponding to ORF M8, 57 and the inter-genic probes in between the 2 ORFs were considered. There were 5 probes complementary to ORF M8, 2 for the inter-genic region and 6 for ORF 57. The

mean signal for each probe was calculated from the 5 replicates on the array, and plotted as a bar chart starting with the most 5' probe (Figure 6.9). The probes corresponding to the inter-genic region between ORFs M8 and 57 have hybridised to complementary cDNA target, which suggests that this region is transcribed. This adds evidence to the prediction that the 2 ORFs are in fact exons of the same gene. Further studies are now required to examine this further and should also examine the abundance of pre-spliced and post-spliced RNAs.

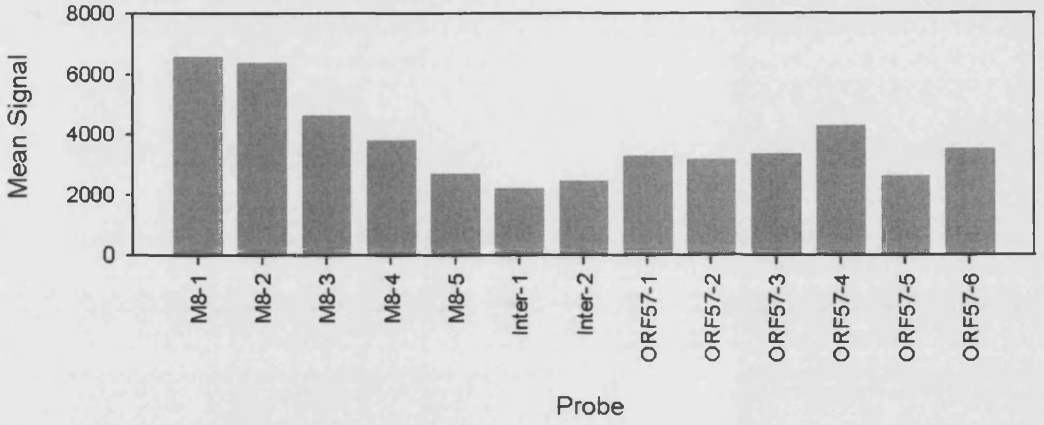


Figure 6.9. Signal strengths of individual probes on the oligo array.
RNA was isolated from NIH 3T3 cells 12h pi by MHV-68. Random nonomers were used to prime labelling of cDNA target, which was hybridised to arrays. The mean signals for probes in the region of ORFs M8 and 57 are shown above.

The high signals observed for probes M8-1 and M8-2 were unexpected, as the remaining M8 probes show less signal. However examining the sequence data for this region shows that there is unlikely to be a polyA signal between M8 and ORF 56, which lies directly upstream of it (Milligan, 1998). Therefore, it is possible that RNA polymerase carries on past the end of ORF 56, and continues some distance into ORF M8. Of course further studies are required to confirm this hypothesis.

6.5 Discussion

While the data produced by the oligonucleotide array should be considered preliminary and further optimisation is required for this system, several conclusions arose from these experiments. Firstly, oligo arrays for MHV-68 were successfully designed and manufactured with the aid of the array group at Arrow Therapeutics. To perform the actual oligo array experiments, the methods developed for the DNA

array system were combined with the established protocols from the array group. Analysing the results of the oligo array experiments showed that overall there was good concordance between the data from the membrane- and glass-based arrays. This provided further confirmation of the results from the membrane array hybridisations, and showed the oligo array system to be functional. There were some differences observed between the two datasets but this was to be expected as different RNA templates were used for the 2 sets of array experiments.

The oligonucleotide array hybridisations produced significantly more data than the previous experiments with the DNA array, which allowed analysis of the viral transcriptome at a much higher resolution than had been possible previously. There are a total of 1438 viral probes, resulting in a 20-fold increase in data for viral transcript abundance. Furthermore each spot was replicated 5 times, which increased the statistical significance of the data produced.

The high resolution of the oligo array allows it to be used for drafting transcript maps. The example shown above examines the abundance of RNA species complementary to the region of the MHV-68 genome around ORFs M8 and 57, as they have been predicted to be exons of the same gene. The oligo array analysis provides experimental evidence for the region between these 2 ORFs being transcribed, and therefore corroborates the exon prediction. The oligo array will most likely not be a practical tool for transcript mapping when used alone, but in collaboration with other tools, the oligo array could provide a powerful method of scanning the entire genome in a single experiment. Furthermore, a strength of this oligo array system is the flexibility of its probeset. As modifying the probeset is fast and simple, the probes can be changed to allow higher resolution data of specific areas of interest on the genome. It is even possible to "walk" along the genome, 1bp at a time by using overlapping probes. The limits to the resolution of such an analysis remain to be seen as it is not clear yet how sensitive the probes on the array are when differentiating between very small differences in hybridisation efficiency. Another important factor when considering the array for transcript mapping, is the intrinsic limitations of array technology. Although the oligo array greatly improves the accuracy of array data, it does not alter the fundamentals of the technology. For example, it is not possible to be certain that the signal for a single probe results from hybridisation of a single target cDNA. If more than 1 cDNA is complementary to the probe sequence, then they will both bind. This can make interpretation of array data more complicated in cases such as overlapping transcripts. However interpretation of such events is only aided by the increased resolution of the oligo array.

As well as increased resolution, the oligo array appears to present increased sensitivity, as demonstrated by the detection of the tRNA sequences that are expressed by MHV-68 (Bowden, 1997), and by lower levels of background. This represents the ability to detect smaller transcripts, and lower transcript levels respectively. Both these features increase the capability of the oligo array. The increased sensitivity in particular could prove to be crucial for analysis of samples in which viral RNAs are rare such as tissue samples from infected animals. Attempts have been made to investigate *in vivo* infection of mice by MHV-68 using the DNA array system. Infected lung and spleen samples were homogenised and RNA extracted for analysis with the array (data not shown). However these experiments proved unsuccessful as the proportion of viral RNA in the samples was below the sensitivity of the DNA array. It would be very interesting to use the oligo array to analyse infected tissue samples.

The oligo array presents a technological advance over the DNA array. The preliminary data presented here confirms this and presents exciting prospects for future applications of the array. In particular, the increased resolution and sensitivity greatly increase the potential of the system for further study of MHV-68 gene expression. While limitations are inherent to any system, this implementation of array technology helps lessen these restrictions, while enhancing its capabilities. Indeed the main hindrance to further applications may be development of tools to analyse and interpret the data produced.

7 Final Discussion

As originally noted in the draft publication of the human genome, "the more we learn about the human genome, the more there is to explore" (Lander, 2001), and this applies equally to the MHV-68 genome. To begin with, the MHV-68 genome has been sequenced twice (Virgin, 1997; Milligan, 1998) and there are a number of differences in the results of the subsequent analyses, with respect to potential ORFs. For example, the analysis performed by Virgin *et al.* (1997) found 2 MHV-68 unique genes (ORFs M5 and M6) in the region located between bp 26,000 and 29,000 of the genome. However, the analysis performed by Milligan *et al.* (1998) did not find either of these, but instead found a separate gene (ORF 17.5) in this same region. ORF 17.5 is homologous to a known herpesvirus scaffolding protein (HSV-1 U_L26.5). One simple method of resolving this type of inconsistency between the 2 sequence analyses would be to use a MHV-68 array, which would provide experimental data for transcription from this region of the genome. Therefore, while sequence analysis is essential for furthering our understanding of the virus, further experimental studies are also required to further elucidate features of the virus.

So far, only a small number of MHV-68 genes have been characterised experimentally, such as the secreted chemokine binding protein encoded by ORF M3 (Parry, 2000; Bridgeman, 2001) and MHV-68 gp150 encoded by ORF M7 (Stewart, 1996). Some wider analyses of gene expression have also been attempted, focusing on a number of genes or a specific region of the genome (Mackett, 1997; Husain, 1999; Rochford, 2001). However, there is much less known of MHV-68's biology than that of other herpesviruses such as HSV-1, EBV and even the relatively recently discovered HHV-8. As MHV-68 is used as an expedient model for other gammaherpesviruses (Simas, 1998b), and has shown itself to be of interest in its own right (Parry, 2000), it is important to characterise the virus further. Therefore, the aim of this thesis was to characterise MHV-68 gene expression through the lytic stage of its life cycle, using techniques that allow comprehensive analysis of gene expression across the entire genome, as opposed to narrow analyses focused on 1 or a few genes.

Array technology (Schena, 1995) is particularly appropriate to use here as so few of MHV-68's genes have been characterised; a global array analysis can provide a complete picture of MHV-68 gene expression, in a quick and relatively simple manner. Studies with similar goals had been successfully performed for other herpesviruses prior to this study, and following this study, there have also been

similar studies of MHV-68 (Stingley, 2000; Paulose-Murphy, 2001; Jenner, 2001; Ebrahimi, 2003; Martinez-Guzman, 2003).

Arrays have also been used to examine the host response to viral infection in terms of gene expression by analysing infected samples with host gene arrays (Zhu, 1998; Chambers, 1999; Simmen, 2001; Tsavachidou, 2001; Carter, 2002; Jones, 2003). So far, host gene arrays do not represent complete gene sets, while producing much more complex datasets due to the larger genomes of the host organisms. This has not prevented their successful application, for example, to identify prognostic markers for various cancers (Dhanasekaran, 2001; van 't Veer, 2002). In terms of viral infections, they have also been successful for identifying genes that are affected by infections. However a greater depth of insight, such as the underlying pathways behind these changes in single gene expressions, has eluded investigators so far.

As the genesets for host organisms are large and incomplete, differential display provides an alternative technique that can overcome these obstacles, and highlight changes in host gene expression (Liang, 1992). In this study, the optimisation of a differential display system for the analysis of changes in gene expression (Chapter 3) has shown that this technique is most suited for highlighting small changes in gene expression, as the redundancy of the dataset increases with the size of the dataset. As such, it has been successfully applied to studies of interferon-responsive RNAs following HCMV infection (Zhu, 1997), changes in host gene expression as latent HSV-1 is stimulated to reactivate (Tal-Singer, 1998) and the stress responses of Salmonella (Wong, 1994). Although not attempted here, differential display could be a very useful technique to aid the study of the changes in host gene expression that occur at the very start of infection. It would be interesting to see if it is possible to dissect between active responses of the host and the effects of host protein synthesis shut-off by the virus, which appears to be very effective during MHV-68 infection (as shown in Chapter 5).

Relatively little is known of the early events during a primary gammaherpesvirus infection. Therefore, CX inhibition of *de novo* viral protein synthesis was used to identify MHV-68 α genes. One α gene, ORF 73, was identified with the MHV-68 DNA array (5.2). It remains to be seen whether MHV-68 encodes more α genes, which were not detected by this array. RT-PCR of RNA samples from CX-inhibited infections showed that faint bands for ORFs K3, 50 and 57, therefore further study of MHV-68 α gene expression is required. It would be interesting to use the more sensitive oligonucleotide array as see if more than 1 α gene is detected.

There have been 2 further array analyses of MHV-68 gene expression since the work presented in this thesis (Ebrahimi, 2003; Martinez-Guzman, 2003). Both studies attempted to elucidate MHV-68 α gene expression in a similar fashion as detailed here in Chapter 5. However Ebrahimi *et al.* (2003) and Martinez-Guzman *et al.* (2003) infected murine epithelial C127 and BHK cells, respectively (murine epithelial NIH 3T3 cells were used here). They also used 100μg/ml and 200μg/ml CX, respectively, to inhibit viral protein synthesis (compared to 2μg/ml here). This suggests that there is a large variability in the potency of CX stocks and/or sensitivity of cell lines to the drug. However, neither of the 2 studies reported that they had measured the efficacy of *de novo* protein synthesis inhibition with their particular cell lines and CX stocks. Ebrahimi *et al.* (2003) reported the detection of 6 α genes (ORFs M4, K3, 38, 50, 57 and 73), of which ORFs K3, 38 and 57 were most highly expressed. Martinez-Guzman *et al.* (2003) detected around 50 genes with CX inhibition (among these the highest were ORFs 69, 73, 74 and 75a), but most of these showed reduced expression relative to levels without CX inhibition. ORFs 38 and 75a are predicted to encode tegument proteins, which suggests that further optimisation of these systems could be useful. However both studies detected ORF 73, in agreement with this study.

These results of studies performed by these 2 MHV-68 arrays highlight some of the difficulties of optimising an array system. While it can be relatively simple to optimise conditions for a single probe for a single gene, it is far more complicated to achieve optimisation across and entire array of probes. Ebrahimi *et al.* (2003) detected levels of ORF 38 transcripts that were equal to or higher than putative α genes, when infection was inhibited by CX. ORF 38 is homologue to a tegument protein of other gammaherpesviruses. In MHV-68 it is predicted to be part of a co-terminal set of genes (ORFs 34-38) sharing a poly-adenylation signal and transcribed together (Rudge, H., personal communication, Milligan, 1998). The homologue of ORF 38 in HSV-1 encodes the U_L11 protein, which accumulates on the cytoplasmic face of host internal membranes during infection, and is thought to play a role in nucleocapsid envelopment and egress (Loomis, 2001). This suggests that ORF 38 is not an α gene, and therefore the signal for ORF 38 observed by Ebrahimi *et al.* (2003) is likely to be noise. As this was one of the highest signals on the array, it brings the validity of the other signals into question. This demonstrates the importance of optimising the whole array so that every probe produces reliable data.

Martinez-Guzman *et al.* (2003) developed a MHV-68 array that showed higher levels of background hybridisation under similar experimental conditions. In fact there was a range of background signals across the array, differing for each probe. As background levels were different for each probe, it would not be valid to compare the signals for one probe against the signals for another probe. For example, 2 probes could show the same signal strength, but this could be due to 90% and 10% background, respectively. This reduces the value of the data from the array as each signal is then only relative to itself. Again this system requires further optimisation to allow use of the array as a single component rather than several individual components.

The MHV-68 DNA array developed here detects only 1 transcript under similar experimental conditions (5.2). Therefore, it does not appear to possess the problems illustrated for the 2 other MHV-68 arrays. However, this array may lack the sensitivity required to detect all α genes. The MHV-68 oligo array that has also been developed here has increased sensitivity over the DNA array. It would therefore be interesting to perform the same experiments with the oligo array.

During uninhibited infection, several genes were observed to be expressed at the same time as ORF 73. These genes were grouped together by cluster analysis and included genes involved in DNA replication, such as DNA polymerase and ribonucleotide reductase. Similar observations have been made for other herpesviruses and subsequently for MHV-68 as well (Paulose-Murphy, 2001; Jenner, 2001; Martinez-Guzman, 2003). This indicates that MHV-68 DNA replication proteins are among the first genes to be expressed during infection. Also the transcript levels of this cluster fall sharply after they peak, which suggests that expression of these genes is inhibited at later stages of infection. It is unclear how this relates to protein levels for those genes, as the two are not necessarily proportional, but it seems likely that protein levels will fall as well.

Clustering the gene expression profiles from an uninhibited infection shows 4 groups, the first being the one just discussed. The second cluster peaks around the same time as the first, but then shows a more sustained expression, and consists predominantly of β genes. The third cluster peaks at a later time point, and later still is the peak for the fourth cluster. The genes in these clusters generally represent γ-1 and γ-2 genes, respectively. However these are general trends and the clusters of transcription profiles are not equivalent to kinetic classes, which is unsurprising as kinetic classes were derived originally from protein abundances (Honess, 1974). Cluster analysis of HHV-8 gene expression following induced reactivation from

latently-infected cell lines shows a similar pattern (Paulose-Murphy, 2001; Jenner, 2001). The lytic gene transcription profiles fall into 3 clusters (Jenner, 2001), and functionally related genes also group together as shown for MHV-68 (5.1.8).

Having performed the first comprehensive analysis of gammaherpesvirus transcription through a *de novo* productive infection here, it is now necessary to build on this foundation to further characterise individual genes and develop the knowledge of MHV-68 biology.

Both the MHV-68 DNA and oligo array systems have much potential for studying the virus further. For example, having characterised lytic infection, latent infection could also be characterised using an array. There is currently only one *in vitro* model for MHV-68 latency, which is the S11E cell line (Usherwood, 1996a). This cell line was isolated from MHV-68 infected mice with lymphoproliferative disease, and found to contain episomal and linear forms of the MHV-68 genome. A proportion of cells in any S11E cell population are known to be lytically infected, making it a less than ideal model.

Preliminary experiments to examine latency in S11E cells were performed using the MHV-68 DNA array developed here. However the results showed that there was a consistent level of lytic cycle-associated transcripts, which prevented the isolation and observation of any potential latent transcripts (data not shown). It was possible to block lytic gene expression by maintaining the cells in medium containing the antiviral 4-'S-EtdU, but array analysis of these genes detected no transcripts (data not shown). Interestingly, northern blot analysis of the S11 cell line, using restriction fragments of the MHV-68 genome, detected only a few transcripts during latency, of which only M2 was abundantly expressed (Husain, 1999). A subsequent array analysis of the S11E cell line detected 1 predominantly expressed gene, ORF 73 (Martinez-Guzman, 2003). It would be interesting to examine the S11 and S11E cell lines further using the sensitive oligo array system.

Studying latent gene expression was also attempted using an *in vivo* model. Preliminary studies of tissue samples from infected mice were performed with the DNA array. However, the MHV-68 DNA array was not sensitive enough to detect viral transcripts in tissue samples (data not shown). Preliminary attempts were made to amplify the RNA template isolated from tissue samples using antisense RNA amplification (Phillips, 1996) with little success (data not shown). However further optimisation of this signal amplification technique is required before its potential can be fully assessed. There are also other methods to produce an increased template population. For example, tissue samples could be dissociated

into a single cell suspension and then sorted by FACS to enrich for infected cells. This would have the effect of increasing the proportion of infected cells in a sample, therefore increasing the amount of viral target in the overall target population. This technique could be particularly useful for enriching infected B cells from spleen samples, as a protocol for this procedure already exists (Marques, 2003). The differential display system could also prove valuable for this type of study, as it would be well suited to highlighting the changes in the expression of a small number of genes. Furthermore, differential display could also be able to highlight changes in host gene expression that occur during MHV-68 latent infection.

An alternative to amplifying the target for the array would be to increase the sensitivity of the probes on the array. The MHV-68 oligo array appears to be more sensitive than the DNA array. Therefore, it would be interesting to use this array to analyse tissue samples. Perhaps combining methods for amplifying target with the increased sensitivity of the oligo array would produce the best results. As well as latent gene expression, *in vivo* lytic gene expression could also be profiled using the oligo array.

A strength of the array system is that it provides a rapid screen of global gene expression. This would be useful, for example, when characterising mutant strains of the virus. The array would show exactly which genes were affected by the mutation, and would therefore be very useful if mutant strains were being created with the purpose of knocking out single genes. A preliminary screen of an ORF M3 deletion mutant (a kind gift from Dr S. Efstathiou, Cambridge University, UK) was performed with the DNA array, which suggested that the transcription of other MHV-68 genes were unaffected by the deletion of ORF M3 (data not shown). The array proved to be a very rapid and convenient method for screening the effects of the mutation on the expression of MHV-68's other genes.

In summary, 2 systems for the analysis of gene expression during MHV-68 infection have been developed and optimised. The differential display system, was found to have an advantage in highlighting changes in both viral and host gene expression. The array system, however, has an advantage in performing a global analysis of gene expression (for the virus only in this study) in a relatively simple manner. The array has been used to characterise the changing MHV-68 transcriptome through a *de novo* lytic infection *in vitro*. This was the first such characterisation for a gammaherpesvirus. A second array system using oligonucleotide probes based on a glass slide was developed and tested in collaboration with the Array Group at Arrow Therapeutics. This system corroborated the results gained with the DNA,

membrane-based array. The oligonucleotide array appeared to be more sensitive than the DNA array and also has the advantage off a flexible design and manufacture stage, which allows the probeset to be modified simply between one array and the next. These features make the oligonucleotide array a more capable system than the DNA array. However, the DNA array technology is far more accessible as its costs are far less than the oligonucleotide array.

As the 2 systems have been established, it would be relatively simple to initiate further studies of MHV-68 biology to further exploit the power of these techniques. In particular the latent life cycle and profiling MHV-68 gene expression *in vivo* would make ideal subjects for continued investigations.

8 Bibliography

Bibliography

Ablashi D., Chatlynne L., Cooper H., Thomas D., Yadav M., Norhanom A. W., Chandana A. K., Churdboonchart V., Kulpradist S. A., Patnaik M., Liegmann K., Masood R., Reitz M., Cleghorn F., Manns A., Levine P. H., Rabkin C., Biggar R., Jensen F., Gill P., Jack N., Edwards J., Whitman J., & Boshoff C. (1999). Seroprevalence of human herpesvirus-8 (HHV-8) in countries of Southeast Asia compared to the USA, the Caribbean and Africa. Br J Cancer **81**:893-7.

Ablashi D. V., Chatlynne L. G., Whitman J. E., Jr., & Cesarman E. (2002). Spectrum of Kaposi's sarcoma-associated herpesvirus, or human herpesvirus 8, diseases. Clin Microbiol Rev **15**:439-64.

Adams M. D., Kerlavage A. R., Fleischmann R. D., Fuldner R. A., Bult C. J., Lee N. H., Kirkness E. F., Weinstock K. G., Gocayne J. D., White O., & et al. (1995). Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature **377**:3-174.

Addison C., Rixon F. J., Palfreyman J. W., O'Hara M., & Preston V. G. (1984). Characterisation of a herpes simplex virus type 1 mutant which has a temperature-sensitive defect in penetration of cells and assembly of capsids. Virology **138**:246-259.

Adler H., Messerle M., Wagner M., & Koszinowski U. H. (2000). Cloning and mutagenesis of the murine gammaherpesvirus 68 genome as an infectious bacterial artificial chromosome. J Virol **74**:6964-74.

Ahuja S. K. and Murphy P. M. (1993). Molecular piracy of mammalian interleukin-8 receptor type B by herpesvirus saimiri. J Biol Chem **268**:20691-20694.

Akula S. M., Wang F. Z., Vieira J., & Chandran B. (2001). Human herpesvirus 8 interaction with target cells involves heparan sulfate. Virology **282**:245-55.

Alba M. M., Das R., Orengo C. A., & Kellam P. (2001). Genomewide function conservation and phylogeny in the Herpesviridae. Genome Res **11**:43-54.

Alber D. G., Powell K. L., Vallance P., Goodwin D. A., & Grahame-Clarke C. (2000). Herpesvirus infection accelerates atherosclerosis in the apolipoprotein E-deficient mouse. Circulation **102**:779-785.

Albrecht J. C., Nicholas J., Biller D., Cameron K. R., Biesinger B., Newman C., Wittmann S., Craxton M. A., Coleman H., Fleckenstein B., & et al. (1992). Primary structure of the herpesvirus saimiri genome. J Virol **66**:5047-58.

Alexander J. M., Nelson C. A., van Berkel V., Lau E. K., Studts J. M., Brett T. J., Speck S. H., Handel T. M., Virgin H. W., & Fremont D. H. (2002). Structural basis of chemokine sequestration by a herpesvirus decoy receptor. Cell **111**:343-56.

Alizadeh A. A., Eisen M. B., Davis R. E., Ma C., Lossos I. S., Rosenwald A., Boldrick J. C., Sabet H., Tran T., Yu X., Powell J. I., Yang L., Marti G. E., Moore T., Hudson J., Jr., Lu L., Lewis D. B., Tibshirani R., Sherlock G., Chan W. C., Greiner T. C., Weisenburger D. D., Armitage J. O., Warnke R., Staudt L. M., & et al. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature **403**:503-11.

Allday M. J. and Crawford D. H. (1988). Role of epithelium in EBV persistence and pathogenesis of B-cell tumours. Lancet **1**:855-7.

Altschul S. F., Grish W., Miller W., Myers E. W., & Lipman D. J. (1990). Basic local alignment search tool. J Mol Biol **215**:403-10.

Ambroziak J. A., Blackbourn D. J., Herndier B. G., Glogau R. G., Gullett J. H., McDonald A. R., Lennette E. T., & Levy J. A. (1995). Herpes-like sequences in HIV-infected and uninfected Kaposi's sarcoma patients. Science **268**:582-3.

Andrewes C. H. (1954). Nomenclature of Viruses. Nature **173**:260-261.

Andrewes C. H., Burnet F. M., Enders J. F., Gard S., Hirst G. K., Kaplan N. M., & Zhdanov V. M. (1961). Taxonomy of viruses infecting vertebrates: present knowledge and ingnorance. Virology **15**:53-5.

Andrews C. H. and Carmichael E. A. (1930). A note on the presence of antibodies to herpesvirus in post-encephalitic and other human sera. Lancet **1**:857-858.

Aoki Y. and Tosato G. (1999). Role of vascular endothelial growth factor/vascular permeability factor in the pathogenesis of Kaposi's sarcoma-associated herpesvirus-infected primary effusion lymphomas. Blood **94**:4247-54.

Arvanitakis L., Geras-Raaka E., Varma A., Gershengorn M. C., & Cesarman E. (1997). Human herpesvirus KSHV encodes a constitutively active G-protein-coupled receptor linked to cell proliferation. Nature **385**:347-350.

Baer R., Bankier A. T., Biggin M. D., Deininger P. L., Farrell P. J., Gibson T. J., Hatfull G., Hudson G. S., Satchwell S. C., Seguin C., & et al. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature **310**:207-11.

Barnes A., Dyson H., Sunil-Chandra N. P., Collins P., & Nash A. A. (1999). 2'-Deoxy-5-ethyl-beta-4'-thiouridine inhibits replication of murine gammaherpesvirus and delays the onset of virus latency. Antivir Chem Chemother **10**:321-6.

Batterson W. and Roizman B. (1983). Characterization of the herpes simplex virion-associated factor responsible for the induction of alpha genes. J Virol **46**:371-7.

Bauer D., Muller H., Reich J., Riedel H., Ahrenkiel V., Warthoe P., & Strauss M. (1993). Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). Nucleic Acids Res **21**:4272-80.

Bellows D. S., Chau B. N., Lee P., Lazebnik Y., Burns W. H., & Hardwick J. M. (2000). Antiapoptotic herpesvirus Bcl-2 homologs escape caspase-mediated conversion to proapoptotic proteins. J Virol **74**:5024-31.

Bigger J. E. and Martin D. W. (2002). Herpesvirus papio 2 encodes a virion host shutoff function. Virology **304**:33-43.

Blackbourn D. J., Lennette E., Klencke B., Moses A., Chandran B., Weinstein M., Glogau R. G., Witte M. H., Way D. L., Kutzkey T., Herndier B., & Levy J. A. (2000). The restricted cellular host range of human herpesvirus 8. Aids **14**:1123-33.

Blasdell K., McCracken C., Morris A., Nash A., Begon M., Bennett M., & Stewart J. (2003). The wood mouse is a natural host for *Murid herpesvirus 4*. J Gen Virol **84**:111-113.

Blaskovic D., Stancekova M., Svobodova J., & Mistrikova J. (1980). Isolation of five strains of herpesviruses from two species of free living small rodents. Acta Virol **24**:468.

Blaskovic D., Stanekova J., & Rajcani J. (1984). Experimental pathogenesis o fmurine herpesvirus in newborn mice. Acta Virol **28**:225-231.

Bloom D. C., Devi-Rao G. B., Hill J. M., Stevens J. G., & Wagner E. K. (1994). Molecular analysis of herpes simplex virus type 1 during epinephrine-induced reactivation of latently infected rabbits in vivo. J Virol **68**:1283-92.

Bookstein R., Lai C. C., To H., & Lee W. H. (1990). PCR-based detection of a polymorphic BamHI site in intron 1 of the human retinoblastoma (RB) gene. Nucleic Acids Res **18**:1666.

Bornkamm G. W. and Hammerschmidt W. (2001). Molecular virology of Epstein-Barr virus. Philos Trans R Soc Lond B Biol Sci **356**:437-59.

Boutolleau D., Fernandez C., Andre E., Imbert-Marcille B. M., Milpied N., Agut H., & Gautheret-Dejean A. (2003). Human herpesvirus (HHV)-6 and HHV-7: two closely related viruses with different infection profiles in stem cell transplantation recipients. J. Infect. Dis. **187**:179-186.

Bowden R. J., Simas J. P., Davis A. J., & Efstathiou S. (1997). Murine gammaherpesvirus 68 encodes tRNA-like sequences which are expressed during latency. J Gen Virol **78**:1675-87.

Brenner S. and Horne R. W. (1959). A negative staining method for high resolution electron microscopy of viruses. Biochimica et Biophysica Acta **34**:103-10.

Bridgeman A., Stevenson P. G., Simas J. P., & Efstathiou S. (2001). A secreted chemokine binding protein encoded by murine gammaherpesvirus- 68 is necessary for the establishment of a normal latent load. J Exp Med **194**:301-12.

Brooks M. A., Ali A. N., Turner P. C., & Moyer R. W. (1995). A rabbitpox virus serpin gene controls host range by inhibiting apoptosis in restrictive cells. J Virol **69**:7688-7698.

Buchen-Osmond, C. The Universal Virus Database of the International Committee on Taxonomy of Viruses (ICTVdb). 2002.
Ref Type: Electronic Citation

Burnet F. M., Keogh E. V., & Lush D. (1937). The immunological reactions of filterable viruses. Australian Journal of Experimental Biology and Medical Science **15**:231-368.

Burnet F. M. and Williams S. M. (1939). Herpes simplex: new point of view. Med J Aust **1**:637-640.

Campadelli-Fiume G., Farabegoli F., Di Gaeta S., & Roizman B. (1991). Origin of unenveloped capsids in the cytoplasm of cells infected with herpes simplex virus 1. J Virol **65**:1589-95.

Campadelli-Fiume G., Stirpe D., Boscaro A., Avitabile E., Foa-Tomasi L., Barker D., & Roizman B. (1990). Glycoprotein C-dependent attachment of herpes simplex virus to susceptible cells leading to productive infection. Virology **178**:213-222.

Campbell M. E., Palfreyman J. W., & Preston C. M. (1984). Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription. J Mol Biol **180**:1-19.

Carter K. L., Cahir-McFarland E., & Kieff E. (2002). Epstein-barr virus-induced changes in B-lymphocyte gene expression. J Virol **76**:10427-36.

Cen H., Williams P. A., McWilliams H. P., Breinig M. C., Ho M., & McKnight J. L. (1993). Evidence for restricted Epstein-Barr virus latent gene expression and anti-EBNA antibody response in solid organ transplant recipients with posttransplant lymphoproliferative disorders. Blood **81**:1393-403.

Cesarman E., Moore P. S., Rao P. H., Inghirami G., Knowles D. M., & Chang Y. (1995). In vitro establishment and characterization of two acquired

immunodeficiency syndrome-related lymphoma cell lines (BC-1 and BC-2) containing Kaposi's sarcoma-associated herpesvirus-like (KSHV) DNA sequences. Blood **86**:2708-14.

Cesarman E., Nador R. G., Bai F., Bohenzky R. A., Russo J. J., Moore P. S., Chang Y., & Knowles D. M. (1996). Kaposi's sarcoma-associated herpesvirus contains G protein-coupled receptor and cyclin D homologs which are expressed in Kaposi's sarcoma and malignant lymphoma. J Virol **70**:8218-8223.

Chambers J., Angulo A., Amaratunga D., Guo H., Jiang Y., Wan J. S., Bittner A., Frueh K., Jackson M. R., Peterson P. A., Erlander M. G., & Ghazal P. (1999). DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. J Virol **73**:5757-66.

Chan S. R. and Chandran B. (2000). Characterization of human herpesvirus 8 ORF59 protein (PF-8) and mapping of the processivity and viral DNA polymerase-interacting domains. J. Virol. **74**:10920-10929.

Chan V., Graves D. J., & McKenzie S. E. (1995). The biophysics of DNA hybridization with immobilized oligonucleotide probes. Biophys J **69**:2243-55.

Chang E. H. Y., Nicholas J., Bellows D. S., Hayward G. S., Guo H.-G., Reitz M. S., & Hardwick J. M. (1997). A bcl-2 homolog encoded by Kaposi's sarcoma-associated virus, human herpesvirus 8, inhibits apoptosis but does not heterodimerize with Bax or Bak. Proc Natl Acad Sci U S A **94**:690-694.

Chang Y., Cesarman E., Pessin M. S., Lee F., Culpepper J., Knowles D. M., & Moore P. S. (1994). Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science **266**:1865-9.

Chang Y., Moore P. S., Talbot S. J., Boshoff C. H., Zarkowska T., Godden-Kent D., Paterson H., Weiss R. A., & Mittnacht S. (1996). Cyclin encoded by KS herpesvirus. Nature **382**:410.

Chee M. S., Bankier A. T., Beck S., Bohni R., Brown C. M., Cerny R., Horsnell T., Hutchison C. A. 3., Kouzarides T., Martignetti J. A., & et al. (1990). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr Top Microbiol Immunol **154**:125-69.

Chenchik A., Moqadam F., & Siebert P. (1995). Marathon cDNA amplification: A new method for cloning full-length cDNAs. CLONTECHniques **10**:5-8.

Cheng N., Trus B. L., Belnap D. M., Newcomb W. W., Brown J. C., & Steven A. C. (2002). Handedness of the herpes simplex virus capsid and procapsid. J Virol **76**:7855-9.

Cheung V. G., Morley M., Aguilar F., Massimi A., Kucherlapati R., & Childs G. (1999). Making and reading microarrays. Nat. Genet. **21**:15-19.

Chmielewicz B., Goltz M., Franz T., Bauer C., Brema S., Ellerbrok H., Beckmann S., Rziha H. J., Lahrmann K. H., Romero C., & Ehlers B. (2003). A novel porcine gammaherpesvirus. Virology **308**:317-29.

Christensen J. P. and Doherty P. C. (1999). Quantitative analysis of the acute and long-term CD4(+) T-cell response to a persistent gammaherpesvirus. J Virol **73**:4279-83.

Chua T. P., Smith C. E., Reith R. W., & Williamson J. D. (1990). Inflammatory responses and the generation of chemoattractant activity in cowpox virus-infected tissues. Immunology **69**:202-208.

Ciampor F., Blaskovic D., & Mistrikova J. (1981). Electron microscopy of rabbit embryo fibroblasts infected with herpervirus isolates from Clethrionomys glareolus and Apodemus flavicollis. Acta Virol **25**:101-107.

Clambey E. T., Virgin H. W. t., & Speck S. H. (2000). Disruption of the murine gammaherpesvirus 68 M1 open reading frame leads to enhanced reactivation from latency. J Virol **74**:1973-84.

Clontech (1998). Atlas Mouse cDNA Expression Array. CLONTECHniques **13**:1-3.

Cole N. L. and Grose C. (2003). Membrane fusion mediated by herpesvirus glycoproteins: the paradigm of varicella-zoster virus. Rev Med Virol **13**:207-22.

Coleman H. M., Lima Bd B., Morton V., & Stevenson P. G. (2003). Murine gammaherpesvirus 68 lacking thymidine kinase shows severe attenuation of lytic cycle replication in vivo but still establishes latency. J Virol **77**:2410-7.

Compton T., Nowlin D. M., & Cooper N. R. (1993). Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. Virology **193**:834-41.

Conley A. J., Knipe D. M., Jones P. C., & Roizman B. (1981). Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by in vitro mutagenesis and defective in DNA synthesis and accumulation of gamma polypeptides. J Virol **37**:191-206.

Cox C., Chang S., Karran L., Griffin B., & Wedderburn N. (1996). Persistent Epstein-Barr virus infection in the common marmoset (Callithrix jacchus). J Gen Virol **77 (Pt 6)**:1173-80.

Crawford D. H. (2001). Biology and disease associations of Epstein-Barr virus. Philos Trans R Soc Lond B Biol Sci **356**:461-73.

Crute J. J. and Lehman I. R. (1989a). Herpes simplex-1 DNA polymerase. Identification of an intrinsic 5'----3' exonuclease with ribonuclease H activity. J Biol Chem **264**:19266-70.

Crute J. J., Tsurumi T., Zhu L. A., Weller S. K., Olivo P. D., Challberg M. D., Mocarski E. S., & Lehman I. R. (1989b). Herpes simplex virus 1 helicase-primase: a complex of three herpes-encoded gene products. Proc Natl Acad Sci U S A **86**:2186-9.

Dalla-Favera R., Bregni M., Erikson J., Patterson D., Gallo R. C., & Croce C. M. (1982). Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci U S A **79**:7824-7.

Davi F., Delecluse H. J., Guiet P., Gabarre J., Fayon A., Gentilhomme O., Felman P., Bayle C., Berger F., Audouin J., Bryon P. A., Diebold J., & Raphael M. (1998). Burkitt-like lymphomas in AIDS patients: characterization within a series of 103 human immunodeficiency virus-associated non-Hodgkin's lymphomas. Burkitt's Lymphoma Study Group. J Clin Oncol **16**:3788-95.

Davison A. J., Dolan A., Akter P., Addison C., Dargan D. J., Alcendor D. J., McGeoch D. J., & Hayward G. S. (2003). The human cytomegalovirus genome

revisited: comparison with the chimpanzee cytomegalovirus genome. J Gen Virol **84**:17-28.

de Jong M. D., Galasso G. J., Gazzard B., Griffiths P. D., Jabs D. A., Kern E. R., & Spector S. A. (1998). Summary of the II International Symposium on Cytomegalovirus. Antiviral Res **39**:141-62.

DeSantis S. M., Pau C. P., Archibald L. K., Nwanyanwu O. C., Kazembe P. N., Dobbie H., Jarvis W. R., & Jason J. (2002). Demographic and immune correlates of human herpesvirus 8 seropositivity in Malawi, Africa. Int J Infect Dis **6**:266-71.

Deshmane S. L. and Fraser N. W. (1989). During latency, herpes simplex virus type 1 DNA is associated with nucleosomes in a chromatin structure. J Virol **63**:943-7.

Dhanasekaran S. M., Barrette T. R., Ghosh D., Shah R., Varambally S., Kurachi K., Pienta K. J., Rubin M. A., & Chinnaiyan A. M. (2001). Delineation of prognostic biomarkers in prostate cancer. Nature **412**:822-6.

Dietz S., Rother K., Bamberger C., Schmale H., Mossner J., & Engeland K. (2002). Differential regulation of transcription and induction of programmed cell death by human p53-family members p63 and p73. FEBS Letters **525**:93-99.

Dingwell K. S., Brunetti C. R., Hendricks R. L., Tang Q., Tang M., Rainbow A. J., & Johnson D. C. (1994). Herpes simplex virus glycoproteins E and I facilitate cell-to-cell spread in vivo and across junctions of cultured cells. J. Virol. **68**:834-845.

Dittmer D., Stoddart C., Renne R., Linquist-Stepps V., Moreno M. E., Bare C., McCune J. M., & Ganem D. (1999). Experimental Transmission of Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) to SCID-hu Thy/Liv mice. J Exp Med **190**:1857-1868.

Domachowske J. B., Bonville C. A., & Rosenberg H. F. (2001). Gene expression in epithelial cells in response to pneumovirus infection. Respir Res **2**:225-33.

Du M. Q., Liu H., Diss T. C., Ye H., Hamoudi R. A., Dupin N., Meignin V., Oksenhendler E., Boshoff C., & Isaacson P. G. (2001). Kaposi sarcoma-associated herpesvirus infects monotypic (IgM lambda) but polyclonal naive B cells

in Castleman disease and associated lymphoproliferative disorders. Blood **97**:2130-6.

Duggan D. J., Bittner M., Chen Y., Meltzer P., & Trent J. M. (1999). Expression profiling using cDNA microarrays. Nat Genet **21**:10-4.

Ebrahimi B., Dutia B. M., Roberts K. L., Garcia-Ramirez J. J., Dickinson P., Stewart J. P., Ghazal P., Roy D. J., & Nash A. A. (2003). Transcriptome profile of murine gammaherpesvirus-68 lytic infection. J Gen Virol **84**:99-109.

Efstathiou S., Ho Y. M., Hall S., Styles C. J., Scott S. D., & Gompels U. A. (1990). Murine herpesvirus 68 is genetically related to the gammaherpesviruses Epstein-Barr virus and herpesvirus saimiri. J Gen Virol **71**:1365-72.

Efstathiou S., Minson A. C., Field H. J., Anderson J. R., & Wildy P. (1986). Detection of herpes simplex virus-specific DNA sequences in latently infected mice and in humans. J Virol **57**:446-55.

Eisen M. B., Spellman P. T., Brown P. O., & Botstein D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A **95**:14863-8.

Epstein A. L., Jacquemont B., & Machuca I. (1984). Infection of a restrictive cell line (XC cells) by intratypic recombinants of HSV-1: relationship between penetration of the virus and relative amounts of glycoprotein C. Virology **132**:315-324.

Epstein M., Achong B., & Barr Y. M. (1957). Virus particles in cultured lymphoblasts from Burkitt's lymphoma. Lancet **1**:702-703.

Epstein M., Achong B., & Barr Y. M. (1965). Morphological and biological studies on a virus in cultured lymphoblasts from Burkitt's lymphoma. J Exp Med **121**:761–70.

Epstein M. A., Morgan A. J., Finerty S., Randle B. J., & Kirkwood J. K. (1985). Protection of cottontop tamarins against Epstein-Barr virus-induced malignant lymphoma by a prototype subunit vaccine. Nature **318**:287-9.

Farrell P. J., Bankier A., Seguin C., Deininger P., & Barrell B. G. (1983). Latent and lytic cycle promoters of Epstein-Barr virus. EMBO J **2**:1331-8.

Faulkner G. C., Burrows S. R., Khanna R., Moss D. J., Bird A. G., & Crawford D. H. (1999). X-Linked agammaglobulinemia patients are not infected with Epstein-Barr virus: implications for the biology of the virus. J Virol **73**:1555-64.

Fenwick M. L. and Everett R. D. (1990). Inactivation of the shutoff gene (UL41) of herpes simplex virus types 1 and 2. J Gen Virol **71 (Pt 12)**:2961-7.

Fenwick M. L. and Walker M. J. (1978). Suppression of the synthesis of cellular macromolecules by herpes simplex virus. J Gen Virol **41**:37-51.

Flint J. and Shenk T. (1997). Viral transactivating proteins. Annu Rev Genet **31**:177-212.

Fodor W. I., Rollins S. A., Bianco-Caron S., Rother R. P., Guilmette E. R., Burton W. V., Albrecht J. C., Fleckenstein B., & Squinto S. P. (1995). The complement control protein homolog of herpesvirus saimiri regulates serum complement by inhibiting C3 convertase activity. J Virol **69**:3889-3892.

Forrester A., Farrell H., Wilkinson G., Kaye J., Davis-Poynter N., & Minson T. (1992). Construction and properties of a mutant of herpes simplex virus type 1 with glycoprotein H coding sequences deleted. J Virol **66**:341-8.

Frenkel N., Schirmer E. C., Wyatt L. S., Katsafanas G., Roffman E., & Danovich R. M. (1990). Isolation of a new herpesvirus from human CD4+ T cells. Proc Natl Acad Sci U S A **87**:748–752.

Friborg J., Jr., Kong W., Hottiger M. O., & Nabel G. J. (1999). p53 inhibition by the LANA protein of KSHV protects against cell death. Nature **402**:889-94.

Gao S. J., Kingsley L., Hoover D. R., Spira T. J., Rinaldo C. R., Saah A., Phair J., Detels R., Parry P., Chang Y., & Moore P. S. (1996a). Seroconversion to antibodies against Kaposi's sarcoma-associated herpesvirus-related latent nuclear antigens before the development of Kaposi's sarcoma. N Engl J Med **335**:233-41.

Gao S. J., Kingsley L., Li M., Zheng W., Parravicini C., Ziegler J., Newton R., Rinaldo C. R., Saah A., Phair J., Detels R., Chang Y., & Moore P. S. (1996b). KSHV antibodies among Americans, Italians and Ugandans with and without Kaposi's sarcoma. Nat Med **2**:925-8.

Garber D. A., Schaffer P. A., & Knipe D. M. (1997). A LAT-associated function reduces productive-cycle gene expression during acute infection of murine sensory neurons with herpes simplex virus type 1. J Virol **71**:5885-93.

Gompels U. A., Craxton M. A., & Honess R. W. (1988). Conservation of gene organization in the lymphotropic herpesviruses herpesvirus Saimiri and Epstein-Barr virus. J Virol **62**:757-67.

Gompels U. A., Nicholas J., Lawrence G., Jones M., Thomson B. J., Martin M. E., Efstathiou S., Craxton M., & Macaulay H. A. (1995). The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. Virology **209**:29-51.

Greshock J., Naylor T. L., Margolin A., Diskin S., Cleaver S. H., Futreal P. A., deJong P. J., Zhao S., Liebman M., & Weber B. L. (2004). 1-Mb resolution array-based comparative genomic hybridization using a BAC clone set optimized for cancer gene analysis. Genome Res. **14**:179-187.

Griffiths P. D. (2002). Tomorrow's challenges for herpesvirus management: potential applications of valacyclovir. J Infect Dis **186 Suppl 1**:131-7.

Grundhoff A. and Ganem D. (2003). The latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus permits replication of terminal repeat-containing plasmids. J Virol **77**:2779-83.

Gutierrez M. I., Bhatia K., Barriga F., Diez B., Muriel F. S., de Andreas M. L., Epelman S., Risueno C., & Magrath I. T. (1992). Molecular epidemiology of Burkitt's lymphoma from South America: differences in breakpoint location and Epstein-Barr virus association from tumors in other world regions. Blood **79**:3261-6.

Hall K. T., Giles M. S., Goodwin D. J., Calderwood M. A., Carr I. M., Stevenson A. J., Markham A. F., & Whitehouse A. (2000). Analysis of gene expression in a human cell line stably transduced with herpesvirus saimiri. J Virol **74**:7331-7.

Hanson N., Henderson G., & Jones C. (1994). The herpes simplex virus type 2 gene which encodes the large subunit of ribonucleotide reductase has unusual regulatory properties. Virus Res **34**:265-80.

Haque M., Chen J., Ueda K., Mori Y., Nakano K., Hirata Y., Kanamori S., Uchiyama Y., Inagi R., Okuno T., & Yamanishi K. (2000). Identification and analysis of the K5 gene of Kaposi's sarcoma- associated herpesvirus. J Virol **74**:2867-75.

Harries A. D., Hargreaves N. J., Chimzizi R., & Salaniponi F. M. (2002). Highly active antiretroviral therapy and tuberculosis control in Africa: synergies and potential. Bull World Health Organ **80**:464-9.

Heller M., van Santen V., & Kieff E. (1982). Simple repeat sequence in Epstein-Barr virus DNA is transcribed in latent and productive infections. J Virol **44**:311-20.

Henderson S., Huen S. D., Rowe M., Dawson C., Johnson G., & Rickinson A. (1993). Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death. Proc Natl Acad Sci U S A **90**:8479-8483.

Hoge A. T., Hendrickson S. B., & Burns W. H. (2000). Murine gammaherpesvirus 68 cyclin D homologue is required for efficient reactivation from latency. J Virol **74**:7016-23.

Honess R. W. and Roizman B. (1973). Proteins specified by herpes simplex virus. XI. Identification and relative molar rates of synthesis of structural and nonstructural herpes virus polypeptides in the infected cell. J Virol **12**:1347-65.

Honess R. W. and Roizman B. (1974). Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J Virol **14**:8-19.

Hughes T. R. and Shoemaker D. D. (2001). DNA microarrays for expression profiling. Current Opinion in Chemical Biology **5**:21-25.

Husain S. M., Usherwood E. J., Dyson H., Coleclough C., Coppola M. A., Woodland D. L., Blackman M. A., Stewart J. P., & Sample J. T. (1999). Murine

gammaherpesvirus M2 gene is latency-associated and its protein a target for CD8(+) T lymphocytes. Proc Natl Acad Sci U S A **96**:7508-13.

Huszar D. and Bacchetti S. (1981). Partial purification and characterization of the ribonucleotide reductase induced by herpes simplex virus infection of mammalian cells. J Virol **37**:580-8.

Hutchinson L., Browne H., Wargent V., Davis-Poynter N., Primorac S., Goldsmith K., Minson A. C., & Johnson D. C. (1992). A novel herpes simplex virus glycoprotein, gL, forms a complex with glycoprotein H (gH) and affects normal folding and surface expression of gH. J Virol **66**:2240-50.

Isegawa Y., Mukai T., Nakano K., Kagawa M., Chen J., Mori Y., Sunagawa T., Kawanishi K., Sashihara J., Hata A., Zou P., Kosuge H., & Yamanishi K. (1999). Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B. J Virol **73**:8053-63.

Jacoby M. A., Virgin H. W. t., & Speck S. H. (2002). Disruption of the M2 gene of murine gammaherpesvirus 68 alters splenic latency following intranasal, but not intraperitoneal, inoculation. J Virol **76**:1790-801.

Jarvis M. A. and Nelson J. A. (2002). Human cytomegalovirus persistence and latency in endothelial cells and macrophages. Curr Opin Microbiol **5**:403-7.

Jenner R. G., Alba M. M., Boshoff C., & Kellam P. (2001). Kaposi's sarcoma-associated herpesvirus latent and lytic gene expression as revealed by DNA arrays. J Virol **75**:891-902.

Jensen K. K., Chen S. C., Hipkin R. W., Wiekowski M. T., Schwarz M. A., Chou C. C., Simas J. P., Alcami A., & Lira S. A. (2003). Disruption of CCL21-induced chemotaxis in vitro and in vivo by M3, a chemokine-binding protein encoded by murine gammaherpesvirus 68. J Virol **77**:624-30.

Jia Q. and Sun R. (2003). Inhibition of gammaherpesvirus replication by RNA interference. J Virol **77**:3301-6.

Johannessen I. and Crawford D. H. (1999). In vivo models for Epstein-Barr virus (EBV)-associated B cell lymphoproliferative disease (BLPD). Rev Med Virol **9**:263-277.

Johnson D. C. and Spear P. G. (1982). Monensin inhibits the processing of herpes simplex virus glycoproteins, their transport to the cell surface, and the egress of virions from infected cells. J Virol **43**:1102-12.

Johnson D. C., Wittels M., & Spear P. G. (1984). Binding to cells of virosomes containing herpes simplex virus type 1 glycoproteins and evidence for fusion. J Virol **52**:238-47.

Jones J. O. and Arvin A. M. (2003). Microarray analysis of host cell gene transcription in response to varicella-zoster virus infection of human T cells and fibroblasts in vitro and SCIDhu skin xenografts in vivo. J Virol **77**:1268-80.

Jones P. C. and Roizman B. (1979). Regulation of herpesvirus macromolecular synthesis. VIII. The transcription program consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. J Virol **31**:299-314.

Jung J. U., Stager M., & Desrosiers R. C. (1994). Virus-encoded cyclin. Mol Cell Biol **14**:7235–7244.

Kapadia S. B., Molina H., van Berkel V., Speck S. H., & Virgin H. W. t. (1999). Murine gammaherpesvirus 68 encodes a functional regulator of complement activation. J Virol **73**:7658-70.

Karlin S., Mocarski E. S., & Schachtel G. A. (1994). Molecular evolution of herpesviruses: genomic and protein sequence comparisons. J Virol **68**:1886-902.

Katz M. H., Hessol N. A., Buchbinder S. P., Hirozawa A., O'Malley P., & Holmberg S. D. (1994). Temporal trends of opportunistic infections and malignancies in homosexual men with AIDS. J Infect Dis **170**:198-202.

Kennedy P. G. (2002). Varicella-zoster virus latency in human ganglia. Rev Med Virol **12**:327-34.

Klemperer H. G., Haynes G. R., Shedden W. I., & Watson D. H. (1967). A virus-specific thymidine kinase in BHK-21 cells infected with herpes simplex virus. Virology **31**:120-8.

Lacoste V., Mauclere P., Dubreuil G., Lewis J., Georges-Courbot M. C., & Gessain A. (2000). KSHV-like herpesviruses in chimps and gorillas. Nature **407**:151-2.

Lacoste V., Mauclere P., Dubreuil G., Lewis J., Georges-Courbot M. C., & Gessain A. (2001). A novel gamma2-herpesvirus of the Rhadinovirus 2 lineage in chimpanzees. Genome Res **11**:1511-1519.

Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J. P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J. C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R. H., Wilson R. K., Hillier L. W., McPherson J. D., Marra M. A., Mardis E. R., Fulton L. A., Chinwalla A. T., Pepin K. H., Gish W. R., Chissoe S. L., Wendl M. C., Delehaunty K. D., Miner T. L., Delehaunty A., Kramer J. B., Cook L. L., Fulton R. S., Johnson D. L., Minx P. J., Clifton S. W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J. F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., Gibbs R. A., Muzny D. M., Scherer S. E., Bouck J. B., Sodergren E. J., Worley K. C., Rives C. M., Gorrell J. H., Metzker M. L., Naylor S. L., Kucherlapati R. S., Nelson D. L., Weinstock G. M., Sakaki Y., Fujiyama A., Hattori M., Yada T., Toyoda A., Itoh T., Kawagoe C., Watanabe H., Totoki Y., Taylor T., Weissenbach J., Heilig R., Saurin W., Artiguenave F., Brottier P., Bruls T., Pelletier E., Robert C., Wincker P., Smith D. R., Doucette-Stamm L., Rubenfield M., Weinstock K., Lee H. M., Dubois J., Rosenthal A., Platzer M., Nyakatura G., Taudien S., Rump A., Yang H., Yu J., Wang J., Huang G., Gu J., Hood L., Rowen L., Madan A., Qin S.,

Davis R. W., Federspiel N. A., Abola A. P., Proctor M. J., Myers R. M., Schmutz J., Dickson M., Grimwood J., Cox D. R., Olson M. V., Kaul R., Shimizu N., Kawasaki K., Minoshima S., Evans G. A., Athanasiou M., Schultz R., Roe B. A., Chen F., Pan H., Ramser J., Lehrach H., Reinhardt R., McCombie W. R., de la Bastide M., Dedhia N., Blocker H., Hornischer K., Nordsiek G., Agarwala R., Aravind L., Bailey J. A., Bateman A., Batzoglou S., Birney E., Bork P., Brown D. G., Burge C. B., Cerutti L., Chen H. C., Church D., Clamp M., Copley R. R., Doerks T., Eddy S. R., Eichler E. E., Furey T. S., Galagan J., Gilbert J. G., Harmon C., Hayashizaki Y., Haussler D., Hermjakob H., Hokamp K., Jang W., Johnson L. S., Jones T. A., Kasif S., Kaspryzk A., Kennedy S., Kent W. J., Kitts P., Koonin E. V., Korf I., Kulp D., Lancet D., Lowe T. M., McLysaght A., Mikkelsen T., Moran J. V., Mulder N., Pollara V. J., Ponting C. P., Schuler G., Schultz J., Slater G., Smit A. F., Stupka E., Szustakowski J., Thierry-Mieg D., Thierry-Mieg J., Wagner L., Wallis J., Wheeler R., Williams A., Wolf Y. I., Wolfe K. H., Yang S. P., Yeh R. F., Collins F., Guyer M. S., Peterson J., Felsenfeld A., Wetterstrand K. A., Patrinos A., Morgan M. J., Szustakowki J., de Jong P., Catanese J. J., Osoegawa K., Shizuya H., Choi S., & Chen Y. J. (2001). Initial sequencing and analysis of the human genome. Nature **409**:860-921.

Lane B. R., Liu J., Bock P. J., Schols D., Coffey M. J., Strieter R. M., Polverini P. J., & Markovitz D. M. (2002). Interleukin-8 and growth-regulated oncogene alpha mediate angiogenesis in Kaposi's sarcoma. J Virol **76**:11570-83.

Lee B. J., Koszinowski U. H., Sarawar S. R., & Adler H. (2003). A gammaherpesvirus G protein-coupled receptor homologue is required for increased viral replication in response to chemokines and efficient reactivation from latency. J Immunol **170**:243-51.

Lee S. I., Murthy S. C. T., rimble J. J., Desrosiers R. C., & Steitz J. A. (1988). Four novel U RNAs are encoded by a herpesvirus. Cell **54**:599-607.

Levine M. and Tjian R. (2003). Transcription regulation and animal diversity. Nature **424**:147-51.

Li M., Lee H., Yoon D. W., Albrecht J. C., Fleckenstein B., Neipel F., & Jung J. U. (1997). Kaposi's sarcoma-associated herpesvirus encodes a functional cyclin. J Virol **71**:1984-1991.

Liang P., Averboukh L., & Pardee A. B. (1993). Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimisation. Nucleic Acids Res **21**:3269-3275.

Liang P. and Pardee A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science **257**:967-71.

Ligas M. W. and Johnson D. C. (1988). A herpes simplex virus mutant in which glycoprotein D sequences are replaced by beta-galactosidase sequences binds to but is unable to penetrate into cells. J Virol **62**:1486-94.

Lipshutz R. J., Fodor S. P., Gingeras T. R., & Lockhart D. J. (1999). High density synthetic oligonucleotide arrays. Nat Genet **21**:20-4.

Liu S., Pavlova I. V., Virgin H. W. t., & Speck S. H. (2000). Characterization of gammaherpesvirus 68 gene 50 transcription. J Virol **74**:2029-37.

Long M. C., Leong V., Schaffer P. A., Spencer C. A., & Rice S. A. (1999). ICP22 and the UL13 protein kinase are both required for herpes simplex virus-induced modification of the large subunit of RNA polymerase II. J Virol **73**:5593-604.

Loomis J. S., Bowzard J. B., Courtney R. J., & Wills J. W. (2001). Intracellular trafficking of the UL11 tegument protein of herpes simplex virus type 1. J. Virol. **75**:12209-12219.

Mackett M., Stewart J. P., de V. P. S., Chee M., Efstathiou S., Nash A. A., & Arrand J. R. (1997). Genetic content and preliminary transcriptional analysis of a representative region of murine gammaherpesvirus 68. J Gen Virol **78**:1425-33.

Magrath I. T. (1991). African Burkitt's lymphoma. History, biology, clinical features, and treatment. Am J Pediatr Hematol Oncol **13**:222-46.

Marchetti M. E., Smith C. A., & Schaffer P. A. (1988). A temperature-sensitive mutation in a herpes simplex virus type 1 gene required for viral DNA synthesis maps to coordinates 0.609 through 0.614 in UL. J Virol **62**:715-21.

Marques S., Efstathiou S., Smith K. G., Haury M., & Simas J. P. (2003). Selective gene expression of latent murine gammaherpesvirus 68 in B lymphocytes. J. Virol. **77**:7308-7318.

Martin S. J., Reutelingsperger C. P., McGahon A. J., Rader J. A., van Schie R. C., LaFace D. M., & Green D. R. (1995). Early redistribution of plasma membrane phosphatidylserine is a general feature of apoptosis regardless of the initiating stimulus: inhibition by overexpression of Bcl-2 and Abl. J Exp Med **182**:1545-56.

Martinez-Guzman D., Rickabaugh T., Wu T. T., Brown H., Cole S., Song M. J., Tong L., & Sun R. (2003). Transcription program of murine gammaherpesvirus 68. J. Virol. **77**:10488-10503.

McGeoch D. J., Cook S., Dolan A., Jamieson F. E., & Telford E. A. (1995). Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. J Mol Biol **247**:443-58.

McGeoch D. J., Dalrymple M. A., Davison A. J., Dolan A., Frame M. C., McNab D., Perry L. J., Scott J. E., & Taylor P. (1988). The complete DNA sequence of the long unique region in the genome of Herpes Simplex virus Type 1. J Gen Virol **69**:1531-1574.

McGeoch D. J., Dolan A., & Ralph A. C. (2000). Toward a comprehensive phylogeny for mammalian and avian herpesviruses. J Virol **74**:10401-6.

Mettenleiter T. C. (2002). Herpesvirus assembly and egress. J Virol **76**:1537-47.

Miller C. L., Lee J. H., Kieff E., & Longnecker R. (1994). An integral membrane protein (LMP2) blocks reactivation of Epstein-Barr virus from latency following surface immunoglobulin crosslinking. Proc Natl Acad Sci U S A **91**:772-6.

Milligan, S., Efstathiou, S., Stewart, J. P., Nash, A. A., and Davison, A. J. Genetic content of murine gammaherpesviruses. 1998.
Ref Type: Electronic Citation

Mistrikova J. and Blaskovic D. (1985). Ecology of the murine alphaherpesvirus and its isolation from lungs of rodents in cell culture. Acta Virol **29**:312-7.

Mistrikova J., Raslova H., Mrmusova M., & Kudelova M. (2000). A murine gammaherpesvirus. Acta Virol **44**:211-26.

Mitsuhashi M., Cooper A., Ogura M., Shinagawa T., Yano K., & Hosokawa T. (1994). Oligonucleotide probe design--a new approach. Nature **367**:759-61.

Miyamoto K. and Morgan C. (1971). Structure and development of viruses as observed in the electron microscope. XI. Entry and uncoating of herpes simplex virus. J Virol **8**:910-8.

Montague M. G. and Hutchinson C. A. (2000). Gene content phylogeny of herpesviruses. Proc Natl Acad Sci U S A **97**:5334-5339.

Muralidhar S., Pumfery A. M., Hassani M., Sadaie M. R., Kishishita M., Brady J. N., Doniger J., Medveczky P., & Rosenthal L. J. (1998). Identification of kaposin (open reading frame K12) as a human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus) transforming gene. J Virol **72**:4980-8.

Murthy S. C., Trimble J. J., & Desrosiers R. C. (1989). Deletion mutants of herpesvirus saimiri define an open reading frame necessary for transformation. J Virol **63**:3307-3314.

Nash A. A. and Sunil-Chandra N. P. (1994). Interactions of the murine gammaherpesvirus with the immune system. Curr Opin Immunol **6**:560-3.

Newcomb W. W., Homa F. L., Thomsen D. R., Trus B. L., Cheng N., Steven A., Booy F., & Brown J. C. (1999). Assembly of the herpes simplex virus procapsid from purified components and identification of small complexes containing the major capsid and scaffolding proteins. J Virol **73**:4239-50.

Nicholas J., Cameron K. R., & Honess R. (1992). Herpesvirus saimiri encodes homologs of G protein-coupled receptors and cyclins. Nature **355**:362-365.

Nicholas J., Smith E. P., Coles L., & Honess R. (1990). Gene expression in cells infected with gammaherpesvirus saimiri: properties of transcripts from two immediate-early genes. Virology **179**:189-200.

Nishioka Y. and Silverstein S. (1977). Degradation of cellular mRNA during infection by herpes simplex virus. Proc Natl Acad Sci U S A **74**:2370-4.

O'Hare P. (1993). The virion transactivator of herpes simplex virus. Seminars in Virology **4**:145-155.

Olsen S. J., Chang Y., Moore P. S., Biggar R. J., & Melbye M. (1998). Increasing Kaposi's sarcoma-associated herpesvirus seroprevalence with age in a highly Kaposi's sarcoma endemic region, Zambia in 1985. Aids **12**:1921-5.

Parry B. C., Simas J. P., Smith V. P., Stewart C. A., Minson A. C., Efstathiou S., & Alcami A. (2000). A broad spectrum secreted chemokine binding protein encoded by a herpesvirus. J Exp Med **191**:573-8.

Paulose-Murphy M., Ha N. K., Xiang C., Chen Y., Gillim L., Yarchoan R., Meltzer P., Bittner M., Trent J., & Zeichner S. (2001). Transcription program of human herpesvirus 8 (kaposi's sarcoma- associated herpesvirus). J Virol **75**:4843-53.

Pearson W. R. and Lipman D. J. (1988). Improved Tools for Biological Sequence Analysis. Proc Natl Acad Sci U S A **85**:2444-8.

Phillips J. and Eberwine J. H. (1996). Antisense RNA Amplification: A Linear Amplification Method for Analyzing the mRNA Population from Single Living Cells. Methods **10**:283-8.

Platt G. M., Simpson G. R., Mittnacht S., & Schulz T. F. (1999). Latent nuclear antigen of Kaposi's sarcoma-associated herpesvirus interacts with RING3, a homolog of the Drosophila female sterile homeotic (fsh) gene. J Virol **73**:9789-95.

Ploubidou A. and Way M. (2001). Viral transport and the cytoskeleton. Curr Opin Cell Biol **13**:97-105.

Pomp D. and Medrano J. F. (1991). Organic solvents as facilitators of polymerase chain reaction. Biotechniques **10**:58-9.

Post L. E. and Roizman B. (1981). A generalized technique for deletion of specific genes in large genomes: alpha gene 22 of herpes simplex virus 1 is not essential for growth. Cell **25**:227-32.

Preston C. M. (2000). Repression of viral transcription during herpes simplex virus latency. J Gen Virol **81**:1-19.

Preston C. M., Frame M. C., & Campbell M. E. (1988). A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. Cell **52**:425-34.

Purifoy D. J., Lewis R. B., & Powell K. L. (1977). Identification of the herpes simplex virus DNA polymerase gene. Nature **269**:621-3.

Quackenbush J. (2001). Computational analysis of microarray data. Nat Rev Genet **2**:418-27.

Rainbow L., Platt G. M., Simpson G. R., Sarid R., Gao S. J., Stoiber H., Herrington C. S., Moore P. S., & Schulz T. F. (1997). The 222- to 234-kilodalton latent nuclear protein (LNA) of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) is encoded by orf73 and is a component of the latency-associated nuclear antigen. J Virol **71**:5915-21.

Rajcani J., Bustamante de Contreras L. R., & Svobodova J. (1986). Corneal infection of murine heroesvirus in mice: the absence of neuronal spread. Acta Virol **31**:25-30.

Ralph D., McClelland M., & Welsh J. (1993). RNA fingerprinting using arbitrarily primed PCR identifies differentially regulated RNAs in mink lung (Mv1Lu) cells growth arrested by transforming growth factor beta 1. Proc Natl Acad Sci U S A **90**:10710-4.

Read G. S. and Frenkel N. (1983). Herpes simplex virus mutants defective in the virion-associated shutoff of host polypeptide synthesis and exhibiting abnormal synthesis of alpha (immediate early) viral polypeptides. J Virol **46**:498-512.

Regenmortel, M. H. V. v., C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and

R. B. Wickner. 2000. Virus Taxonomy: The Classification and Nomenclature of Viruses. The Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.

Renne R., Blackbourn D., Whitby D., Levy J., & Ganem D. (1998). Limited transmission of Kaposi's sarcoma-associated herpesvirus in cultured cells. J Virol **72**:5182-8.

Renne R., Zhong W., Herndier B., McGrath M., Abbey N., Kedes D., & Ganem D. (1996). Lytic growth of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in culture. Nat Med **2**:342-6.

Reschke, M. Glycoprotein B (gpUL55) mediated morphogenesis of infectious HCMV particles. 1995. Biografix.
Ref Type: Electronic Citation

Revello M. G. and Gerna G. (2002). Diagnosis and management of human cytomegalovirus infection in the mother, fetus, and newborn infant. Clin Microbiol Rev **15**:680-715.

Robinson J. and Miller G. (1975). Assay for Epstein-Barr virus based on stimulation of DNA synthesis in mixed leukocytes from human umbilical cord blood. J Virol **15**:1065-1072.

Rochford R., Lutzke M. L., Alfinito R. S., Clavo A., & Cardin R. D. (2001). Kinetics of murine gammaherpesvirus 68 gene expression following infection of murine cells in culture and in mice. J Virol **75**:4955-63.

Rochford R. and Mosier D. E. (1995). Differential Epstein-Barr virus gene expression in B-cell subsets recovered from lymphomas in SCID mice after transplantation of human peripheral blood lymphocytes. J Virol **69**:150-5.

Roizman B. (1996). Herpesviridiae - an update. Arch Virol **123**:425-49.

Roizman, B. and D. M. Knipe. 2001a. Herpes Simplex Viruses and Their Replication, p. 2399. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (eds.), Fields Virology. Lippincott Williams & Wilkins.

Roizman B. and Whitley R. J. (2001b). The nine ages of herpes simplex virus. Herpes **8**:23-7.

Rose T. M., Strand K. B., Schultz E. R., Schaefer G., Rankin G. W., Jr., Thouless M. E., Tsai C. C., & Bosch M. L. (1997). Identification of two homologs of the Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in retroperitoneal fibromatosis of different macaque species. J Virol **71**:4138-44.

Rowe W. P., Hartley J. W., & Waterman S. (1956). Cytopathogenic agent resembling human salivary gland virus disease. Proc Soc Exp Biol Med **92**:418-424.

Rozen, S. and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers, p. 365-86. *In* S. Misener (ed.), Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ.

Rubin R. H. (2001). Cytomegalovirus in solid organ transplantation. Transpl. Infect. Dis. **3 Suppl 2**:1-5.

Russo J. J., Bohenzky R. A., Chien M. C., Chen J., Yan M., Maddalena D., Parry J. P., Peruzzi D., Edelman I. S., Chang Y., & Moore P. S. (1996). Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). Proc Natl Acad Sci U S A **93**:14862-7.

Salahuddin S. Z., Ablashi D. V., Markham P. D., Josephs S. F., Sturzenegger S., & Kaplan M. (1986). Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. Science **234**:596–601.

Sander, D. The Big Picture Book of Viruses: Herpesviridae. 2003. 2003.
Ref Type: Electronic Citation

Sarmiento M., Haffey M., & Spear P. G. (1979). Membrane proteins specified by herpes simplex viruses. III. Role of glycoprotein VP7(B2) in virion infectivity. J Virol **29**:1149-58.

Savard M., Belanger C., Tardif M., Gourde P., Flamand L., & Gosselin J. (2000). Infection of primary human monocytes by Epstein-Barr virus. J Virol **74**:2612-2619.

Schaffer P. A., Aron G. M., Biswal N., & Benyesh-Melnick M. (1973). Temperature-sensitive mutants of herpes simplex virus type 1: isolation, complementation and partial characterization. Virology **52**:57-71.

Schek N. and Bachenheimer S. L. (1985). Degradation of cellular mRNAs induced by a virion-associated factor during herpes simplex virus infection of Vero cells. J Virol **55**:601-10.

Schena M., Shalon D., Davis R. W., & Brown P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270**:467-70.

Sciortino M. T., Suzuki M., Taddeo B., & Roizman B. (2001). RNAs extracted from herpes simplex virus 1 virions: apparent selectivity of viral but not cellular rnas packaged in virions. J Virol **75**:8105-16.

Secchiero P., Sun D., De Vico A. L., Crowley R. W., Reitz M. S., Jr., Zauli G., Lusso P., & Gallo R. C. (1997). Role of the extracellular domain of human herpesvirus 7 glycoprotein B in virus binding to cell surface heparan sulfate proteoglycans. J Virol **71**:4571-80.

Sharp P. M. (2002). Origins of human virus diversity. Cell **108**:305-12.

Sheaffer A. K., Newcomb W. W., Gao M., Yu D., Weller S. K., Brown J. C., & Tenney D. J. (2001). Herpes simplex virus DNA cleavage and packaging proteins associate with the procapsid prior to its maturation. J Virol **75**:687-98.

Sheldrick P. and Berthelot N. (1975). Inverted repetitions in the chromosome of herpes simplex virus. Cold Spring Harb Symp Quant Biol **39 Pt 2**:667-78.

Shimizu T., Kawakita S., Li Q. H., Fukuhara S., & Fujisawa J. (2003). Human T-cell leukemia virus type 1 Tax protein stimulates the interferon-responsive enhancer element via NF-kappaB activity. FEBS Lett **539**:73-7.

Shukla D. and Spear P. G. (2001). Herpesviruses and heparan sulfate: an intimate relationship in aid of viral entry. J Clin Invest **108**:503-10.

Simas J. P., Bowden R. J., Paige V., & Efstathiou S. (1998a). Four tRNA-like sequences and a serpin homologue encoded by murine gammaherpesvirus 68 are dispensable for lytic replication in vitro and latency in vivo. J Gen Virol **79**:149-53.

Simas J. P. and Efstathiou S. (1998b). Murine gammaherpesvirus 68: a model for the study of gammaherpesvirus pathogenesis. Trends Microbiol **6**:276-82.

Simas J. P., Swann D., Bowden R., & Efstathiou S. (1999). Analysis of murine gammaherpesvirus-68 transcription during lytic and latent infection. J Gen Virol **80**:75-82.

Simmen K. A., Singh J., Luukkonen B. G., Lopper M., Bittner A., Miller N. E., Jackson M. R., Compton T., & Fruh K. (2001). Global modulation of cellular transcription by human cytomegalovirus is initiated by viral glycoprotein B. Proc Natl Acad Sci U S A **98**:7140-5.

Simmons A. (2001). Herpesvirus and multiple sclerosis. Herpes **8**:60-3.

Sinclair A. J., Palmero I., Holder A., Peters G., & Farrell P. J. (1995). Expression of cyclin D2 in Epstein-Barr virus positive Burkitt's lymphoma cell lines is related to methylation status of the gene. J Virol **69**:1292-1295.

Sissons J. G., Bain M., & Wills M. R. (2002). Latency and reactivation of human cytomegalovirus. J Infect **44**:73-7.

Skepper J. N., Whiteley A., Browne H., & Minson A. (2001). Herpes simplex virus nucleocapsids mature to progeny virions by an envelopment --> deenvelopment --> reenvelopment pathway. J. Virol. **75**:5697-5702.

Smibert C. A., Johnson D. C., & Smiley J. R. (1992). Identification and characterization of the virion-induced host shutoff product of herpes simplex virus gene UL41. J Gen Virol **73 (Pt 2)**:467-70.

Smith J. S. and Robinson N. J. (2002). Age-specific prevalence of infection with herpes simplex virus types 2 and 1: a global review. J Infect Dis **186 Suppl 1**:3-28.

Smith M. G. (1956). Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus disease. Proc Soc Exp Biol Med **92**:424-430.

Smith T. F. and Waterman M. S. (1981). Identification of common molecular subsequences. J Mol Biol **147**:195-7.

Sodeik B. (2000). Mechanisms of viral transport in the cytoplasm. Trends Microbiol **8**:465-72.

Sodeik B., Ebersold M. W., & Helenius A. (1997). Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. J Cell Biol **136**:1007-21.

Southern E., Mir K., & Shchepinov M. (1999). Molecular interactions on microarrays. Nat Genet **21**:5-9.

Spear P. G. and Roizman B. (1972). Proteins specified by herpes simplex virus. V. Purification and structural proteins of the herpesvirion. J Virol **9**:143-59.

Speck P., Haan K. M., & Longnecker R. (2000). Epstein-Barr virus entry into cells. Virology **277**:1-5.

Speck S. H., Chatila T., & Flemington E. (1997). Reactivation of Epstein-Barr virus: regulation and function of the BZLF1 gene. Trends Microbiol **5**:399-405.

Spellman P. T., Sherlock G., Zhang M. Q., Iyer V. R., Anders K., Eisen M. B., Brown P. O., Botstein D., & Futcher B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell **9**:3273-97.

Spiller O. B., Robinson M., O'Donnell E., Milligan S., Morgan B. P., Davison A. J., & Blackbourn D. J. (2003). Complement regulation by Kaposi's sarcoma-associated herpesvirus ORF4 protein. J Virol **77**:592-9.

Spivack J. G. and Fraser N. W. (1987). Detection of herpes simplex virus type 1 transcripts during latent infection in mice. J Virol **61**:3841-7.

Stevens J. G., Wagner E. K., Devi-Rao G. B., Cook M. L., & Feldman L. T. (1987). RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. Science **235**:1056-9.

Stevenson P. G., May J. S., Smith X. G., Marques S., Adler H., Koszinowski U. H., Simas J. P., & Efstathiou S. (2002). K3-mediated evasion of CD8(+) T cells aids amplification of a latent gamma-herpesvirus. Nat Immunol **3**:733-40.

Stewart J. P., Janjua N. J., Pepper S. D., Bennion G., Mackett M., Allen T., Nash A. A., & Arrand J. R. (1996). Identification and characterization of murine gammaherpesvirus 68 gp150: a virion membrane glycoprotein. J Virol **70**:3528-35.

Stewart J. P., Micali N., Usherwood E. J., Bonina L., & Nash A. A. (1999). Murine gamma-herpesvirus 68 glycoprotein 150 protects against virus-induced mononucleosis: A model system for gamma- herpesvirus vaccination. Vaccine **17**:152-157.

Stewart J. P., Usherwood E. J., Ross A., Dyson H., & Nash T. (1998). Lung epithelial cells are a major site of murine gammaherpesvirus persistence. J Exp Med **187**:1941-51.

Stingley S. W., Ramirez J. J., Aguilar S. A., Simmen K., Sandri-Goldin R. M., Ghazal P., & Wagner E. K. (2000). Global analysis of herpes simplex virus type 1 transcription using an oligonucleotide-based DNA microarray. J Virol **74**:9916-27.

Sugden B. (1994). Latent infection of B lymphocytes by Epstein-Barr virus. Sem Virol **5**:197-205.

Sun R., Lin S. F., Staskus K., Gradoville L., Grogan E., Haase A., & Miller G. (1999). Kinetics of Kaposi's sarcoma-associated herpesvirus gene expression. J Virol **73**:2232-42.

Sunil-Chandra N. P., Arno J., Fazakerley J., & Nash A. A. (1994). Lymphoproliferative disease in mice infected with murine gammaherpesvirus 68. Am J Pathol **145**:818-26.

Sunil-Chandra N. P., Efstathiou S., Arno J., & Nash A. A. (1992a). Virological and Pathological Features of Mice Infected with Murine Gammaherpesvirus-68. J Gen Virol **73**:2347-2356.

Sunil-Chandra N. P., Efstathiou S., & Nash A. A. (1992b). Murine gammaherpesvirus 68 establishes a latent infection in mouse B lymphocytes in vivo. J Gen Virol **73**:3275-9.

Sunil-Chandra N. P., Efstathiou S., & Nash A. A. (1993). Interactions of murine gammaherpesvirus 68 with B and T cell lines. Virology **193**:825-33.

Svobodova J., Blaskovic D., & Mistrikova J. (1982a). Growth characteristics of herpesviruses isolated from free living small rodents. Acta Virol **26**:256-63.

Svobodova J., Stancekova M., Blaskovic D., Mistrikova J., Lesso J., Russ G., & Masarova P. (1982b). Antigenic relatedness of alphaherpesviruses isolated from free-living rodents. Acta Virol **26**:438-43.

Swain M. A. and Galloway D. A. (1986). Herpes simplex virus specifies two subunits of ribonucleotide reductase encoded by 3'-coterminal transcripts. J Virol **57**:802-8.

Swaminathan S., Huneycutt B. S., Reiss C. S., & Kieff E. (1992). Epstein-Barr virus-encoded small RNAs (EBERs) do not modulate interferon effects in infected lymphocytes. J Virol **66**:5133-5136.

Swaminathan S., Tomkinson B., & Kieff E. (1991). Recombinant Epstein-Barr virus with small RNA (EBER) genes deleted transforms lymphocytes and replicates in vitro. Proc Natl Acad Sci U S A **88**:1546-1550.

Tal-Singer R., Podrzucki W., Lasner T. M., Skokotas A., Leary J. J., Fraser N. W., & Berger S. L. (1998). Use of differential display reverse transcription-PCR to reveal cellular changes during stimuli that result in herpes simplex virus type 1 reactivation from latency: upregulation of immediate-early cellular response genes TIS7, interferon, and interferon regulatory factor-1. J Virol **72**:1252-61.

Talbot S. J., Weiss R. A., Kellam P., & Boshoff C. (1999). Transcriptional analysis of human herpesvirus-8 open reading frames 71, 72, 73, K14, and 74 in a primary effusion lymphoma cell line. Virology **257**:84-94.

Taub R., Kirsch I., Morton C., Lenoir G., Swan D., Tronick S., Aaronson S., & Leder P. (1982). Translocation of the c-myc gene into the immunoglobulin heavy

chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc Natl Acad Sci U S A **79**:7837-41.

Telford E. A., Watson M. S., Aird H. C., Perry J., & Davison A. J. (1995). The DNA sequence of equine herpesvirus 2. J Mol Biol **249**:520-8.

Thomas S. K., Lilley C. E., Latchman D. S., & Coffin R. S. (2002). A protein encoded by the herpes simplex virus (HSV) type 1 2-kilobase latency-associated transcript is phosphorylated, localized to the nucleus, and overcomes the repression of expression from exogenous promoters when inserted into the quiescent HSV genome. J Virol **76**:4056-67.

Trus B. L., Heymann J. B., Nealon K., Cheng N., Newcomb W. W., Brown J. C., Kedes D. H., & Steven A. C. (2001). Capsid structure of Kaposi's sarcoma-associated herpesvirus, a gammaherpesvirus, compared to those of an alphaherpesvirus, herpes simplex virus type 1, and a betaherpesvirus, cytomegalovirus. J Virol **75**:2879-90.

Tsavachidou D., Podrzucki W., Seykora J., & Berger S. L. (2001). Gene array analysis reveals changes in peripheral nervous system gene expression following stimuli that result in reactivation of latent herpes simplex virus type 1: induction of transcription factor Bcl-3. J Virol **75**:9909-17.

Usherwood E. J., Stewart J. P., & Nash A. A. (1996a). Characterization of tumor cell lines derived from murine gammaherpesvirus-68-infected mice. J Virol **70**:6516-8.

Usherwood E. J., Stewart J. P., Robertson K., Allen D. J., & Nash A. A. (1996b). Absence of splenic latency in murine gammaherpesvirus 68-infected B cell-deficient mice. J Gen Virol **77**:2819-25.

van 't Veer L. J., Dai H., van de Vijver M. J., He Y. D., Hart A. A., Mao M., Peterse H. L., van der Kooy K., Marton M. J., Witteveen A. T., Schreiber G. J., Kerkhoven R. M., Roberts C., Linsley P. S., Bernards R., & Friend S. H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature **415**:530-6.

van Berkel V., Barrett J., Tiffany H. L., Fremont D. H., Murphy P. M., McFadden G., Speck S. H., & Virgin H. I. (2000). Identification of a gammaherpesvirus selective chemokine binding protein that inhibits chemokine action. J Virol **74**:6741-7.

van Berkel V., Preiter K., Virgin H. W. t., & Speck S. H. (1999). Identification and initial characterization of the murine gammaherpesvirus 68 gene M3, encoding an abundantly secreted protein. J Virol **73**:4524-9.

van Dyk L. F., Hess J. L., Katz J. D., Jacoby M., Speck S. H., & Virgin H. I. (1999). The murine gammaherpesvirus 68 v-cyclin gene is an oncogene that promotes cell cycle progression in primary lymphocytes. J Virol **73**:5110-22.

van Santen V. L. (1991). Characterization of the bovine herpesvirus 4 major immediate-early transcript. J Virol **65**:5211-5224.

Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G., Smith H. O., Yandell M., Evans C. A., Holt R. A., Gocayne J. D., Amanatides P., Ballew R. M., Huson D. H., Wortman J. R., Zhang Q., Kodira C. D., Zheng X. H., Chen L., Skupski M., Subramanian G., Thomas P. D., Zhang J., Gabor Miklos G. L., Nelson C., Broder S., Clark A. G., Nadeau J., McKusick V. A., Zinder N., Levine A. J., Roberts R. J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z., Di Francesco V., Dunn P., Eilbeck K., Evangelista C., Gabrielian A. E., Gan W., Ge W., Gong F., Gu Z., Guan P., Heiman T. J., Higgins M. E., Ji R. R., Ke Z., Ketchum K. A., Lai Z., Lei Y., Li Z., Li J., Liang Y., Lin X., Lu F., Merkulov G. V., Milshina N., Moore H. M., Naik A. K., Narayan V. A., Neelam B., Nusskern D., Rusch D. B., Salzberg S., Shao W., Shue B., Sun J., Wang Z., Wang A., Wang X., Wang J., Wei M., Wides R., Xiao C., Yan C., Yao A., Ye J., Zhan M., Zhang W., Zhang H., Zhao Q., Zheng L., Zhong F., Zhong W., Zhu S., Zhao S., Gilbert D., Baumhueter S., Spier G., Carter C., Cravchik A., Woodage T., Ali F., An H., Awe A., Baldwin D., Baden H., Barnstead M., Barrow I., Beeson K., Busam D., Carver A., Center A., Cheng M. L., Curry L., Danaher S., Davenport L., Desilets R., Dietz S., Dodson K., Doup L., Ferriera S., Garg N., Gluecksmann A., Hart B., Haynes J., Haynes C., Heiner C., Hladun S., Hostin D., Houck J., Howland

T., Ibegwam C., Johnson J., Kalush F., Kline L., Koduru S., Love A., Mann F., May D., McCawley S., McIntosh T., McMullen I., Moy M., Moy L., Murphy B., Nelson K., Pfannkoch C., Pratts E., Puri V., Qureshi H., Reardon M., Rodriguez R., Rogers Y. H., Romblad D., Ruhfel B., Scott R., Sitter C., Smallwood M., Stewart E., Strong R., Suh E., Thomas R., Tint N. N., Tse S., Vech C., Wang G., Wetter J., Williams S., Williams M., Windsor S., Winn-Deen E., Wolfe K., Zaveri J., Zaveri K., Abril J. F., Guigo R., Campbell M. J., Sjolander K. V., Karlak B., Kejariwal A., Mi H., Lazareva B., Hatton T., Narechania A., Diemer K., Muruganujan A., Guo N., Sato S., Bafna V., Istrail S., Lippert R., Schwartz R., Walenz B., Yooseph S., Allen D., Basu A., Baxendale J., Blick L., Caminha M., Carnes-Stine J., Caulk P., Chiang Y. H., Coyne M., Dahlke C., Mays A., Dombroski M., Donnelly M., Ely D., Esparham S., Fosler C., Gire H., Glanowski S., Glasser K., Glodek A., Gorokhov M., Graham K., Gropman B., Harris M., Heil J., Henderson S., Hoover J., Jennings D., Jordan C., Jordan J., Kasha J., Kagan L., Kraft C., Levitsky A., Lewis M., Liu X., Lopez J., Ma D., Majoros W., McDaniel J., Murphy S., Newman M., Nguyen T., Nguyen N., & Nodell M. (2001). The sequence of the human genome. Science **291**:1304-51.

Virgin H. W. t., Latreille P., Wamsley P., Hallsworth K., Weck K. E., Dal Canto A. J., & Speck S. H. (1997). Complete sequence and genomic analysis of murine gammaherpesvirus 68. J Virol **71**:5894-904.

Virgin H. W. t., Presti R. M., Li X. Y., Liu C., & Speck S. H. (1999). Three distinct regions of the murine gammaherpesvirus 68 genome are transcriptionally active in latently infected mice. J Virol **73**:2321-32.

Wagner E. K. and Bloom D. C. (1997). Experimental investigation of herpes simplex virus latency. Clin Microbiol Rev **10**:419-43.

Weller S. K., Aschman D. P., Sacks W. R., Coen D. M., & Schaffer P. A. (1983). Genetic analysis of temperature-sensitive mutants of HSV-1: the combined use of complementation and physical mapping for cistron assignment. Virology **130**:290-305.

Weller T. H. and Stoddard M. B. (1952). Intranuclear intrusion bodies in cultures of human tissue innoculated with varicella vesicle fluid. J Immunol **68**:311-319.

Whitby D., Stossel A., Gamache C., Papin J., Bosch M., Smith A., Kedes D. H., White G., Kennedy R., & Dittmer D. P. (2003). Novel Kaposi's sarcoma-associated herpesvirus homolog in baboons. J Virol **77**:8159-65.

Whitehouse A. (2003). Herpesvirus saimiri: a potential gene delivery vector (review). Int. J. Mol. Med. **11**:139-148.

Whiteley A., Bruun B., Minson T., & Browne H. (1999). Effects of targeting herpes simplex virus type 1 gD to the endoplasmic reticulum and trans-Golgi network. J Virol **73**:9515-20.

Whitley, R. J. 2001. Herpes Simplex Viruses, p. 2461. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (eds.), Fields Virology. Lippincott Williams & Wilkins.

Whitley R. J., Kimberlin D. W., & Roizman B. (1998). Herpes simplex viruses. Clin Infect Dis **26**:541-53.

Wildy P. and Watson D. H. (1963). Electron microscopic studies on the architecture of animal viruses. Quant Biol **27**:25-47.

Wong K. K. and McClelland M. (1994). Stress-inducible gene of Salmonella typhimurium identified by arbitrarily primed PCR of RNA. Proc Natl Acad Sci U S A **91**:639-43.

Wu C. A., Nelson N. J., McGeoch D. J., & Challberg M. D. (1988). Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis. J Virol **62**:435-43.

Wu L., Lo P., Yu X., Stoops J. K., Forghani B., & Zhou Z. H. (2000a). Three-dimensional structure of the human herpesvirus 8 capsid. J Virol **74**:9646-54.

Wu T. T., Usherwood E. J., Stewart J. P., Nash A. A., & Sun R. (2000b). Rta of murine gammaherpesvirus 68 reactivates the complete lytic cycle from latency. J Virol **74**:3659-67.

WuDunn D. and Spear P. G. (1989). Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. J Virol **63**:52-8.

Wymer J. P., Aprhys C. M., Chung T. D., Feng C. P., Kulka M., & Aurelian L. (1992). Immediate early and functional AP-1 cis-response elements are involved in the transcriptional regulation of the large subunit of herpes simplex virus type 2 ribonucleotide reductase (ICP10). Virus Res **23**:253-70.

Xiao P. and Capone J. P. (1990). A cellular factor binds to the herpes simplex virus type 1 transactivator Vmw65 and is required for Vmw65-dependent protein-DNA complex assembly with Oct-1. Mol Cell Biol **10**:4974-7.

Yang W. C., Devi-Rao G. V., Ghazal P., Wagner E. K., & Triezenberg S. J. (2002). General and specific alterations in programming of global viral gene expression during infection by VP16 activation-deficient mutants of herpes simplex virus type 1. J Virol **76**:12758-74.

York I. A., Roop C., Andrews D. W., Riddell S. R., Graham F. L., & Johnson D. C. (1994). A cytosolic herpes simplex virus protein inhibits antigen presentation to CD8+ T lymphocytes. Cell **77**:525-35.

Zalani S., Holley-Guthrie E., & Kenney S. (1996). Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism. Proc Natl Acad Sci U S A **93**:9194-9.

Zamora R., Alarcon L., Vodovotz Y., Betten B., Kim P. K., Gibson K. F., & Billiar T. R. (2001). Nitric oxide suppresses the expression of Bcl-2 binding protein BNIP3 in hepatocytes. J Biol Chem **276**:46887-95.

Zhu F. X., Cusano T., & Yuan Y. (1999). Identification of the immediate-early transcripts of Kaposi's sarcoma-associated herpesvirus. J Virol **73**:5556-67.

Zhu H., Cong J. P., Mamtora G., Gingeras T., & Shenk T. (1998). Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. Proc Natl Acad Sci U S A **95**:14470-5.

Zhu H., Cong J. P., & Shenk T. (1997). Use of differential display analysis to assess the effect of human cytomegalovirus infection on the accumulation of cellular RNAs: induction of interferon-responsive RNAs. Proc Natl Acad Sci U S A **94**:13985-90.

Ziegler J. L. and Katongole-Mbidde E. (1996). Kaposi's sarcoma in childhood: an analysis of 100 cases from Uganda and relationship to HIV infection. Int J Cancer **65**:200-203.

9 Appendices

9.1 Appendix I

9.1.1 DNA Array Primer Sequences

Primer	5' Sequence	3' Sequence
A1 - M1	gccaaagcatagctcactgg	acaccttgatgacccttcc
A2 - M2	caaggaaagattcccaatcc	caggacttggtacaggactcg
A3 - M3	gctcattaaatgctgcatcc	tctcttgacccagctcttcc
A4 - M4	cacccagcctagatttttag	tacgcagacataatagctgc
A5 - ORF4	gggattgtgggtgtaaatgg	gggatcacagtaaaacactgc
A6 - ORF6	ctcccaggctcctcttgg	ccacgtccatgaaaaattgg
A7 - ORF7	gattatgcatggacctcagc	caccatagaagcaggcaagg
A8 - ORF8	atgagagtcgcccacctaac	tgaagatgtagggcacgatg
A9 - ORF9	cagccttgccaagaaagagc	ctgaacggaatgtcctcgc
A10 - ORF10	ttaatcactggtgcgtgacc	cgggagattagtttttatagcc
A11 - ORF11	acgaaggtcatctacatttgc	tgtgaggacaacctggagg
A12 - K3	aggagagttctgttggatctgc	tccccatcactatcatcagc
B1 - M5	agcatacagcgtgggaatgg	aggggatttccaggtagagg
B2 - M6	ggaccctccattctataaacc	ccagctcgggaggggggc
B3 - ORF17	cgtgggaggatatgtggac	atgtgtgcaacatctgcagc
B4 - ORF18	aggtgcatcagccatgttgg	tgaggtccttcgttgtcagg
B5 - ORF19	tttgtttacctgaacccaaacc	actccacccgactggaagg
B6 - ORF20	gaatgaattaggagccaagc	cttctctggtggtactgtgc
B7 - ORF21	aacaaccctggattccaacc	gcccatatgcccctatatcc
B8 - ORF22	gttggtgcatggggtagg	gctggtaatggtgggatagc
B9 - ORF23	aaagactgtgccttaacagg	gtcatggtcaggatttttgc
B10 - ORF24	tgtgaatatgatgagaatagcc	atatcactggaacataggtgc
B11 - ORF25	tttgcctatgtcaggactgg	acttggtcttcttccatcacc
B12 - ORF26	ctctaacatctagattgtatgc	ctaaatcatcatgctcatccc
C1 - ORF27	tggtgaggttgtgctaaatgg	ggcaggataatagggatgtttg
C2 - ORF29b	tgttacagaaggatgccaag	agacttccgtgtgattgc
C3 - ORF30	ggatgctctgtgaataaatcaaag	tgtcttggccgcgcttgc
C4 - ORF31	ccaagacaagtgtatgatgc	acaatgtcaccaaacagtgc
C5 - ORF32	ttcctttacaacaagatgagc	tctaaagaagccctcatccc
C6 - ORF33	gacatttctgaataaagagtgtatatgg	gtactggaatggctgacttcc
C7 - ORF29a	tataccgcagcacagtttcc	agcacccaactagtgttttagc
C8 - ORF34	ttttccccgaagcagatcc	tatgctgtgctcactgatgg
C9 - ORF35	aattgggagtagtttaagggc	ttaacagtgtcttccaactcc
C10 - ORF36	tggattaccgacagttaccg	tctgatgccatggtagaaacc
C11 - ORF37	ggaagggtcgattattctgg	atacacggcagacacctcac
C12 - ORF38	tcggatgttgcaagaaaacc	cgtgtttgcctcgttcatattc

D1 - ORF39	aaaagttggagccgtctagg	cagacggtacactgttactgc
D2 - ORF40	caatggcacgatgttcagg	tctccttgccagaagtaccc
D3 - ORF42	cttcccaggctgatgttgg	ccaccatggaaatcagtgc
D4 - ORF43	tctgtcagagagtgccatgc	aagtgttgctacccaagaagc
D5 - ORF44	acctcagatgccaaaattcg	tgtgggacaatatgacttgagc
D6 - ORF45	atggacccctttaagaaacc	cttgtacttgactcgctgacc
D7 - ORF46	tggacacttggctaaaaacg	tgaacactaaaggccaaacc
D8 - ORF47	tggtccttttgtgtattaatgc	tgggccacattgatattccc
D9 - ORF48	accttgaaacccgtgaagg	gagctggcacacaaagaagg
D10 - ORF49	ccttctctggaaagcgtgg	aattgacagtgcctatggcc
D11 - ORF50	caaagtccataacaggcatcc	aaatgcctcaacttctctgg
D12 - M7	tgtggcgttaaatccctagc	tggggaagttggttctgagg
E1 - ORF52	tggtcaaggaagtagaaagg	ctggcacccacagtagtttc
E2 - ORF53	atcacccaagaaaccacacc	gcgtagatcaaaaagacaccac
E3 - ORF54	atactcctttgtgcccaagc	attggcccggttgaatctcc
E4 - ORF55	tgtaccttacaagaggctcg	gcctcatctacacttattgc
E5 - ORF56	atggccagataccacagc	aaaactgtttggcatcatcg
E6 - M8	accagttgaggagccaacg	tttctggttccatctgtttgg
E7 - ORF57	cttgcctgaaacacggtagg	agagggtccctgattacacg
E8 - ORF58	attgtgggaggaatgtctgc	ggaacccatgtggaaagc
E9 - ORF59	cttccagcttgacactgagc	aagacagggaggcagattcc
E10 - ORF60	ctgggttccttgaactgacc	cttcagcttcatcatgtcacc
E11 - ORF61	acgtggagcctgttcagacc	gctgccagataaccattttcc
E12 - ORF62	tcttcagtagtcacatcagc	aggctcttgttgtgttttagg
F1 - ORF63	ccatcagtgagcgatagtgg	tgcagcacaaagaagactgg
F2 - ORF64	ttcgcatcgaaggtacagc	gccaaacatttcagcagagg
F3 - M9	cccagagctccataacaagc	aaatgctccagaagaggaagg
F4 - ORF66	cagtggatgagtttagaagc	agggcacagtgagtattttgg
F5 - ORF67	ctgatagacgagctctgtgg	gccatattgaccctgttgc
F6 - ORF68	tccttcctctcaaatacactgg	atgcagcacgtagaagcagagg
F7 - ORF69	gcgctcaacaggctctgc	acccagattgtccgtgtgg
F10 - M10c	aagcgaggagcagcacagc	tggtatgtcaacccctgacc
F11 - ORF72	tccaaggatttcttgacagc	atgtaggccctgacctttcc
F12 - M11	tgggcaaccctgattacagc	atgatcctccgtccaactgc
G1 - ORF73	acacaacctcaggcaaacc	ccttcaacatcaacatctgg
G2 - ORF74	cttagaaaactcatcattgtcc	acacaatagcatagatcctgg
G3 - ORF75c	agacagaaaaagaactcatcg	gaatcagatcgtgaagatagg
G4 - ORF75b	ggatgaggacgtctgggc	ataaaatctagccgctgggc
G5 - ORF75a	ttggacattgagggtcttcc	gctttgcaggtgagagtatgc
G6 - M12	ggggaaaatatgcgtgatacc	ccagtggtctgttctgatgg
G7 - M13	tagtagggggcctcctgc	caaagtttaaagtgaaagtaagc
G8 - M14	gctaccgcccgggcctgagg	agcagggcccgagcccctctt
G9 - M1-M2	ctggggaccagatgtaaagc	tccatgggtgcatatttgg
G10 -M2-M1	taacagtgaaggtgctaacg	caggttctcggttcaagtcc
G11 - M2-M3	aaaccccctccagtaaaagg	caaggccccagagaaagc
G12 - M3-M2	ttaaaaaagatacgagtcaggtgg	cggtcgagaagacatatccc

H1 - M3-M4	aaatatgctccatggtttgg	tccccaaaagatacatcaagc
H2 - M4-M3	tcatctgacttgctgcatacg	cttgtgagcctcaagagtgg
H3 - K3-M5	ttttggccaggagaatataagc	cgagacaggttgtggaaagc
H4 - M5-K3	caggagacatggcctatcg	cccctgaagcataactctcg
H5 - ORF27-29b	tctggattttcatttcatgtg	cacgtgacaagagtgcttgg
H6 - ORF29-27	Use same primers as H5	
H7 - M10c-ORF72	tcatggcaacagtcaaaagg	aaacgtgctagccaatcagg
H8 - ORF72-M10c	ctgtgcgagatttgcgtatg	atgtatcgatccccgtcctc
H9 - GAPDH	acccagaagactgtggatgg	acctggtcctcagtgtagcc
H10 - Myosin 1	gaaggtcatgggtgttctgg	gttcagcttttccaggaagg
H11 - MOD	gcgtcactccctttacgc	ggaatcatcagtggcaatcc
H12 - Beta-actin	ggactcctatgtgggtgacg	atcacaatgcctgtggtacg
J1 - Cab45	gatcagcgctaaggagatgc	ggaactcgtcctcagtcagc
J2 - Ribosomal protein S29	agcagctctactggagtcacc	aaagactagcatgatcggttcc
J3 - Ubiquitin	tcttcgtgaagaccctgacc	gttctcgatggtgtcactgg
J4 - Phospholipase A2	ccagtcacagcaagcatacc	cctgcttctgcttcatctcc
J5 - HPRT	gggggctataagttctttgc	ttcgagaggtccttttcacc
J6 - TMV		
J7 - Luciferase	atgcatgcggccgcatctag	ccaacaccggcataaagaattg

9.2 Appendix II

9.2.1 Layout of Vacuum-Spotted Arrays

	1	2	3	4	5	6	7	8	9	10	11	12
A	M1	M2	M3	M4	ORF4	ORF6	ORF7	ORF8	ORF9	ORF10	Cab45	K3
B	M5	M6	ORF17	ORF18	ORF19	ORF2	ORF21	ORF22	ORF23	ORF24	ORF25	ORF26
C	ORF27	ORF29b	ORF30	ORF31	ORF32	ORF33	ORF29a	ORF34	ORF35	ORF36	ORF37	ORF38
D	ORF39	ORF40	ORF42	ORF43	ORF44	ORF45	ORF46	ORF47	ORF48	ORF49	ORF50	M7
E	ORF52	ORF53	ORF54	ORF55	ORF56	M8	ORF57	ORF58	ORF59	ORF60	ORF61	ORF62
F	ORF63	ORF64	M9	ORF66	ORF67	ORF68	ORF69	Ribo S29	Luc	M10c	ORF72	M11
G	ORF73	ORF74	ORF75c	ORF75b	Ph A2	M12	M13	TMV	M1-M2	M2-M1	M2-M3	M3-M2
H	M3-M4	M4-M3	K3-M5	M5-K3	ORF27-29b	HPRT	M10c-ORF72	ORF72-M10c (1/2)	GAPDH	Myosin 1	MOD	β-actin

9.2.2 Layout of Biomek Arrays

	1	2	3	4	5	6	7	8	9	10	11	12
A	**M1**	**M2**	**M3**	**M4**	**ORF4**	**ORF6**	**ORF7**	**ORF8**	**ORF9**	**ORF10**	Cab45t	**K3**
B	**M5**	**M6**	**M6**	**ORF18**	**ORF19**	**ORF20**	**ORF21**	**ORF22**	**ORF23**	**ORF24**	**ORF25**	**M6**
C	**ORF27**	**ORF29b**	**ORF30**	**ORF31**	**ORF32**	**ORF33**	**ORF29a**	**ORF34**	**ORF35**	**ORF36**	**ORF37**	**ORF38**
D	**ORF39**	**ORF40**	**ORF42**	**ORF43**	**ORF44**	**ORF45**	**ORF46**	**ORF47**	**ORF48**	**ORF49**	**ORF50**	**M7**
E	**ORF52**	**ORF53**	**ORF54**	**ORF55**	**ORF56**	**M8**	**ORF57**	**ORF58**	**ORF59**	**ORF60**	**ORF61**	**ORF62**
F	**ORF57**	**ORF57**	**M9**	**ORF66**	**ORF67**	**ORF68**	**ORF69**	Ribo S29	Luc	**M10c**	**ORF72**	**M11**
G	**ORF73**	**ORF74**	**ORF75c**	**ORF75b**	Ph A2	**M12**	**M13**	TMV	M1-M2	M2-M1	M2-M3	M3-M2
H	M3-M4	M4-M3	K3-M5	M5-K3	ORF27-29b	HPRT	M10c-ORF72	ORF72-M10c (1/2)	GAPDH	Myosin 1	MOD	β-actin

9.2.3 Layout of Pin-Tool Arrays

	1	2	3	4	5	6	7	8	9	10	11	12
A	M1	M2	M3	M4	ORF4	ORF6	ORF7	ORF8	ORF9	ORF10	ORF11	K3
B	M5	M6	ORF17	ORF18	ORF19	ORF2	ORF21	ORF22	ORF23	ORF24	ORF25	ORF26
C	ORF27	ORF29b	ORF30	ORF31	ORF32	ORF33	ORF29a	ORF34	ORF35	ORF36	ORF37	ORF38
D	ORF39	ORF40	ORF42	ORF43	ORF44	ORF45	ORF46	ORF47	ORF48	ORF49	ORF50	M7
E	ORF52	ORF53	ORF54	ORF55	ORF56	M8	ORF57	ORF58	ORF59	ORF60	ORF61	ORF62
F	ORF63	ORF64	M9	ORF66	ORF67	ORF68	ORF69	H_2O	Bluescript	M10c	ORF72	M11
G	ORF73	ORF74	ORF75c	ORF75b	H_2O	M12	M13	TMV	M1-M2	M2-M1	M2-M3	M3-M2
H	M3-M4	M4-M3	K3-M5	M5-K3	ORF27-29b	pGEM	M10c-ORF72	ORF72-M10c	GAPDH	Myosin 1	MOD	β-actin
I	Cab45	Ribo S29	Ubiquitin	Ph A2	HPRT	TMV	Luc					

9.3 Appendix III – The State of Knowledge of MHV-68 Genes, Predicted and Confirmed, at the Start of Array Development

Gene	Genome Location	Coding Strand	Function (putative unless confirmed)	Reference / Comments
M1	2023 - 3282	R	Serpin	Disruption enhances reactivation from latency - Clambey et al, 2000, J. Virol. 74:1973-84
M2	4031 - 4627	L		
M3	6060 - 7277	L	Broad spectrum secreted chemokine binding protein	Parry et al, 2000, J.Exp. Med., 191:573-8
M4	8409 - 9785	R		
ORF4	9873 - 11036	R	Complement regulatory protein	
ORF6	11215 - 14523	R	Single-stranded DNA binding protein	
ORF7	14526 - 16481	R	Transport protein	
ORF8	16505 - 19051	R	Glycoprotein B	
ORF9	19217 - 22297	R	DNA polymerase	
ORF10	22269 - 23522	R		
ORF11	23488 - 24651	R		
K3	24733 - 25335	L	BHV4 IE1 homolog	
M5	26178 - 26672	R		

M6	26554 - 28308	R	
ORF17	28326 - 29957	L	Capsid protein
ORF18	29917 - 30768	R	
ORF19	30726 - 32273	L	Tegument protein
ORF20	32119 - 32880	L	
ORF21	32879 - 34810	R	Thymidine kinase
ORF22	34833 - 37022	R	Glycoprotein H
ORF23	37025 - 38167	L	
ORF24	38103 - 40253	L	
ORF25	40263 - 44381	R	Major capsid protein
ORF26	44423 - 45319	R	Capsid protein
ORF27	45329 - 46090	R	
ORF29b	46395 - 47438	L	Packaging protein
ORF30	47507 - 47746	R	
ORF31	47710 - 48309	R	
ORF32	48294 - 49625	R	
ORF33	49588 - 50568	R	
ORF29a	50549 -	L	Packaging protein

	51466		
ORF34	51465 - 52460	R	
ORF35	52423 - 52878	R	
ORF36	52847 - 54157	R	Kinase
ORF37	54129 - 55586	R	Alkaline exonuclease
ORF38	55544 - 55768	R	Myristylated tegument protein
ORF39	55802 - 56950	L	Glycoprotein M
ORF40	57046 - 58875	R	Helicase-primase
ORF42	58876 - 59634	L	
ORF43	59634 - 61334	L	Capsid protein
ORF44	61303 - 63630	R	Helicase-primase
ORF45	63655 - 64272	L	
ORF46	64275 - 65021	L	Uracil DNA glycosylase
ORF47	65027 - 65545	L	Glycoprotein L
ORF48	65584 - 66582	L	
ORF49	66741 - 67643	L	
ORF50	67907 - 69373	R	Transcriptional activator
M7	69466 -	R	Glycoprotein 150

	70914			
ORF52	70960 - 71369	L		
ORF53	71447 - 71701	L	Reactivation of complete lytic cycle from latency	Wu et al, 2000, J. Virol. 74:3659-67
ORF54	71806 - 72702	R	dUTPase	
ORF55	72744 - 73313	L		
ORF56	73289 - 75793	R	DNA replication protein	
M8	76015 - 76485	R		
ORF57	76662 - 77159	R	Immediate-early protein	
ORF58	77214 - 78254	L		
ORF59	78258 - 79439	L	DNA replication protein	
ORF60	79565 - 80479	L	Ribonucleotide reductase, small	
ORF61	80517 - 82865	L	Ribonucleotide reductase, large	
ORF62	82871 - 84010	L	Assembly/DNA maturation	
ORF63	83751 - 86564	R	Tegument protein	
ORF64	86567 - 93937	R	Tegument protein	
M9	93962 - 94519	L		
ORF66	94515 - 95741	L		
ORF67	95738 -	L	Tegument protein	

	96415			
ORF68	96673 - 98052	R	Glycoprotein	
ORF69	98061 - 98936	R		
M10a	98903 - 101224	R		
M10b	99087 - 101204	R		
M10c	99187 - 101367	R		
ORF72	102426 - 103181	L	Cyclin D homolog	
M11	103418 - 103930	R	BCL-2 homolog (gene 16?)	Inhibits Fas- and TNF-induced apoptosis - Wang et al, 1999, J Gen Virol, 80:2737-40. Unlike the Bcl-homologs of other gammaHV's, MHV-68's can be digested by caspases. But the products of digestion are not proapoptiotic - Bellows et al J. Virol. 74:5024-31
ORF73	103927 - 104868	L	Immediate-early protein	
ORF74	105057 - 106067	R	GCR (IL-8 receptor homolog?)	
ORF75c	106070 - 109999	L	Tegument protein/FGARAT	
ORF75b	110077 - 113901	L	Tegument protein/FGARAT	
ORF75a	114032 - 117904	L	Tegument protein/FGARAT	
M12	117992 - 118681	R		
M13	118149 - 118784	R		
M14	118808 - 119125	L		
- M1 - M2	3283 -			

	4030			
M2 - M3	4628 - 6059			
M3 - M4	7278 - 8408			
K3 - M5	25336 - 26177			
ORF27 - 29b	46091 - 46394			
M10c - ORF72	101368 - 102425			